
Digitized by the Internet Archive
in 2022 with funding from
Kahle/Austin Foundation

https://archive.org/details/introductiontoth0002arth

# INTRODUCTION TO THEORETICAL PHYSICS

# INTRODUCTION TO THEORETICAL PHYSICS

530.151
Hlle
tV5
1928
V.2

VOLUME II

BY
ARTHUR HAAS, PH.D.
PROFESSOR OF PHYSICS IN THE UNIVERSITY OF VIENNA

TRANSLATED BY
T. VERSCHOYLE, M.C., PH.D., A.R.C.S.

*SECOND EDITION*

LONDON
CONSTABLE & COMPANY LTD
10–12 ORANGE STREET W.C. 2
1929

PUBLISHED BY

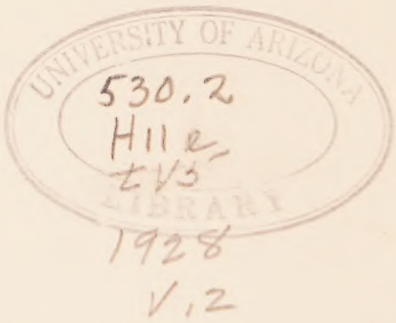
*Constable & Company Limited*
*London W.C. 2*

BOMBAY
CALCUTTA MADRAS
LEIPZIG

*Oxford University*
*Press*

TORONTO

*The Macmillan Company*
*of Canada, Limited*




*Printed in Great Britain by The Whitefriars Press, Ltd.,*
*London and Tonbridge.*

# TRANSLATOR'S NOTE

During the four years which have elapsed since the publication of the first English edition of Volume II., immense strides have been made in the development of atomic theory, which has undergone a complete transformation consequent on the introduction of the wave theory of matter and of quantum mechanics. Part III., which is a translation of the second German edition of the author's "Atomic Theory," has therefore been almost entirely rewritten. Professor Haas has, moreover, supplied several new sections for Part IV. (Theory of Heat). Only minor alterations have been made in Part V.

T. VERSCHOYLE.

*June*, 1929.



86958

# CONTENTS

## PART III

### ATOMIC THEORY

### CHAPTER XI

#### Electrons Atoms and Light Quanta


| | | PAGE |
|---|---|---|
| § 86. | The Elementary Quantum of Electricity | 3 |
| § 87. | Electrolysis, and the Mass of the Hydrogen Atom | 9 |
| § 88. | Negative Radiation and the Electron | 12 |
| § 89. | Positive Radiation | 16 |
| § 90. | The Elementary Quantum of Action | 21 |
| § 91. | Light Corpuscles and Matter Waves | 22 |
| § 92. | The Nuclear Atom | 26 |
| § 93. | Characteristic Potentials and Spectral Terms | 28 |
| § 94. | The Sequence of the Elements | 33 |
| § 95. | Chemical Periodicity | 37 |


### CHAPTER XII

#### The Principles of Atomic Mechanics


| | | |
|---|---|---|
| § 96. | De Broglie's Theory | 41 |
| § 97. | Schrödinger's Theory | 46 |
| § 98. | The Quantum Mechanics of Heisenberg | 51 |
| § 99. | The Statistical Interpretation of Material Waves | 57 |
| § 100. | The Angular Momentum of Electrons and Atoms | 58 |
| § 101. | The Multiplicity of Spectral Terms | 63 |
| § 102. | The Resolution of Spectral Terms in a Magnetic Field | 65 |
| § 103. | Pauli's Principle | 69 |


### CHAPTER XIII

#### Atomic Spectra


| | | |
|---|---|---|
| § 104. | The Spectrum of Hydrogen | 72 |
| § 105. | The Spectrum of Ionized Helium | 78 |
| § 106. | The Spectra of Atoms with One Valency Electron | 84 |

PAGE

§ 107. The Spectra of Atoms with Two Valency Electrons 90
§ 108. The Spectral Terms in Equivalent Systems of Electrons . . . . . 95
§ 109. Spectra of Elements in the Short Periods . . 99
§ 110. Spectra of Elements in the Long Periods . . 101
§ 111. X-Ray Spectra . . . . . 107
§ 112. The Periodic System . . . . . 118
§ 113. The Zeeman Effect . . . . . 125
§ 114. The Stark Effect . . . . . 133
§ 115. Intensities of Spectral Lines . . . . 137
§ 116. The Magneton . . . . . 139
§ 117. The Excitation of Atoms . . . . . 144

## CHAPTER XIV

### X-Rays

§ 118. Space Lattices . . . . . 153
§ 119. Lattice Planes . . . . . 159
§ 120. The Interference of X-Rays in Crystals . . 165
§ 121. The Relation between Electron Rays and X-Rays . 172

## CHAPTER XV

### The Atomic Nucleus

§ 122. Isotopy . . . . . 175
§ 123. The Packing Effect . . . . . 181
§ 124. Disintegration of the Nucleus, and Transformation of the Elements . . . . . 184
§ 125. Radioactive Equilibrium . . . . . 187
§ 126. The Transformation Series . . . . . 193
§ 127. Material Radiation of Radioelements . . . 202
§ 128. Gamma-Radiation and Nuclear Levels . . 207
§ 129. Cosmic Rays . . . . . 211
§ 130. The Disruption of the Elements . . . . 212
§ 131. The Structure of Atomic Nuclei, and Harkins' Rules . . . . . 215

## CHAPTER XVI

### The Molecule

§ 132. Rotation-Vibration Spectra . . . . 220
§ 133. Band Spectra . . . . . 223
§ 134. Nuclear Intervals and Dissociation Energy . . 231
§ 135. Chemical Compounds . . . . . 235
§ 136. The Atomic Lattice . . . . . 240

PAGE
§ 137. Electrolytes . . . . . 242
§ 138. The Polarity of Molecules . . . . . 243
§ 139. Metals . . . . . 244

## CHAPTER XVII

### The Interaction between Light and Matter

§ 140. Dispersion . . . . . 248
§ 141. The Photo-Electric Effect . . . . . 250
§ 142. The Transformation of Electronic Radiation into Light-Quanta . . . . . 252
§ 143. The Compton Effect . . . . . 253
§ 144. Dirac's Equations . . . . . 258
§ 145. The Transformation of Matter into Light . . 261

# PART IV

## THEORY OF HEAT

## CHAPTER XVIII

### Statistics

§ 146. Statistical Probability . . . . . 267
§ 147. The Modulus of Distribution . . . . . 272
§ 148. The Canonical Distribution of Energy . . 276
§ 149. Equilibrium and Modular Equalization . . 279
§ 150. Homogeneous Systems . . . . . 281
§ 151. Cyclic Processes . . . . . 284
§ 152. Theorem of Equipartition . . . . . 288
§ 153. Theorem of the Virial . . . . . 290
§ 154. Rayleigh and Jeans's Principle . . . . 292
§ 155. The Statistical Equilibrium of a Vibrating System 296

## CHAPTER XIX

### Thermodynamics

§ 156. Temperature Scales . . . . . 301
§ 157. Calorimetry . . . . . 304
§ 158. The Two Principal Laws of Thermodynamics . 306
§ 159. Irreversibility . . . . . 308
§ 160. Free Energy and Thermodynamic Potential . 310
§ 161. Changes of State in a Homogeneous System . 312
§ 162. The Joule-Thomson Effect . . . . . 318
§ 163. The Coexistence of Phases . . . . . 321
§ 164. Equilibrium in Heterogeneous Systems . . 327

PAGE

§ 165. The Perfect Gas . . . . . 330
§ 166. Osmotic Pressure . . . . . 339
§ 167. The Equation of State of van der Waals . . 343
§ 168. Law of Mass Action . . . . . 346
§ 169. The Brownian Movement . . . . . 348
§ 170. Sedimentation Equilibria . . . . . 352

## CHAPTER XX

### The Quantum Theory of Heat

§ 171. The Specific Heat of Solid Bodies . . . 355
§ 172. Nernst's Heat Theorem . . . . . 359
§ 173. The Absolute Value of Entropy . . . . 361
§ 174. Cavity Radiation . . . . . 366
§ 175. Planck's Law of Radiation . . . . 369

# PART V

## THEORY OF RELATIVITY

## CHAPTER XXI

### Special Theory of Relativity

§ 176. Michelson's Experiment . . . . . 377
§ 177. Einstein's Principle of Relativity . . . 381
§ 178. The Minkowski-World . . . . . 383
§ 179. The Lorentz-Transformation . . . . 385
§ 180. Relativity of Scales of Time and Length . . 388
§ 181. Composition of Velocities . . . . . 391
§ 182. Fizeau's Experiment . . . . . 394
§ 183. Relativistic Dynamics . . . . . 396
§ 184. The Inertial Mass of Energy . . . . 400
§ 185. World-Tensors . . . . . 401
§ 186. The Electromagnetic Field Tensor . . . 405
§ 187. The Matter Tensor . . . . . 407

## CHAPTER XXII

### Theory of Gravitation

§ 188. The Law of Gravitation as a Generalization of the Law of Inertia . . . . . 409
§ 189. Gaussian Coordinate Systems . . . . . 418
§ 190. Generalized Conception of Tensors . . . 420

PAGE
§ 191. The Metric Fundamental Tensor . . . 423
§ 192. Geodesic Lines . . . . . 426
§ 193. The Riemann Tensor . . . . . 430
§ 194. Einstein's and Newton's Laws of Gravitation . 434
§ 195. The Metrics of Space and Time in a Gravitational Field . . . . . 440
§ 196. The Experimental Confirmation of Einstein's Theory of Gravitation . . . . . 444
§ 197. Size and Mass of the Universe . . . . 445

## APPENDIX

Summary of the Contents of Volume II. . . . . 449
Synopsis of the Symbols most frequently used . . 476
The Universal Constants of Physics . . . . 478
Addenda and Corrigenda . . . . . 479
Name and Subject Indexes to Volume II. . . . 482

# INTRODUCTION TO THEORETICAL PHYSICS

## PART III

## ATOMIC THEORY

## Chapter XI

# ELECTRONS ATOMS AND LIGHT QUANTA

### § 86. The Elementary Quantum of Electricity.

In the whole of theoretical physics there is no older conception than that of the *individual structure of matter*, according to which matter consists of minute, invisible, *individual particles*. This idea originated long ago with the Greek philosopher *Democritus*, who gave the name *atoms* to the ultimate invisible particles into which matter could be resolved. This term has also been retained in modern science for the *smallest particles of the chemical elements*, and it is to the building up of atoms to form so-called molecules that the newer chemistry, founded by *Dalton*, ascribes the multiplicity of chemical compounds.

Numerous physical phenomena, however, have clearly shown that *atoms* must undoubtedly *contain electrical charges*. Above all, the fact of *electrolysis* demonstrates that atoms are influenced by electrical forces. Moreover, after the establishment of the electromagnetic theory of light, the long known phenomenon of the *dispersion of light* effected by matter could not be explained without the assumption of a mutual action between electromagnetic waves and electrical charges contained in the atoms of matter. The hypothesis of an atomistic structure of matter thus necessarily led to the hypothesis of an *atomistic structure of electricity*, and hence to the conception of an *elementary quantum of electrical charge* (or elementary charge). Not only theoretical considerations, however, have given rise to this fundamental assumption of modern atomic theory, for a *direct empirical confirmation* is to be found in some *experiments of Millikan*, which have, moreover, rendered possible a direct and extremely accurate determination of the elementary quantum of electricity.

Millikan's experiments consisted in the observation with a microscope of *minute, individual particles of matter* carrying *feeble electric charges*, and moving under the two-fold influence of gravity and of a field directed vertically upwards.

This method of so-called "*individual observation*" [1] was devised in 1909 by *Ehrenhaft* and a little later also by Millikan, and was further developed and perfected in the following years.[2]

Millikan worked with *oil droplets* with a diameter of about $10^{-4}$ cm. The advantage of oil droplets, as compared with particles of other substances, is that they do not change in size owing to evaporation. By means of a spray-diffuser a fine mist of oil is blown into a chamber, and through a small hole in the bottom of the latter a droplet is from time to time admitted into a condenser placed beneath the chamber, the condenser consisting of two parallel horizontal plates. In consequence of the friction due to the spraying, these droplets are almost always electrically charged. The plates of the condenser can be connected through a commutator to a battery of accumulators giving some 10,000 volts. The particle in the interior of the condenser is illuminated by a strong source of light placed at one side, and it is at the same time observed by a microscope through a window in the condenser; it then appears as a bright point on a black background, although its true shape cannot be recognized. When the electric field is cut out, a particle of this kind slowly falls with a constant velocity, in consequence of its own weight and the strong frictional resistance. When, however, the battery is switched on, the particle slowly rises, again with a constant velocity, provided that the electric force is not much greater than that of gravity. (With the given size of droplet and potential difference the conditions are just right.) By always switching on or off at the correct moment, it is possible to keep a droplet moving up and down between the two plates of the condenser as long as is required.

On account of the large friction, the *velocity* of the particle in both the upward and the downward movement is *proportional to the effective force*, as follows both from mechanical considerations and by direct experimental confirmation. Denoting, therefore, the velocity of the downward motion by $v_1$ and that of the upward motion by $v_2$, the mass of the particle by $m$ and its charge by $Q$, the acceleration due to

[1] The essential novelty of this method lay in the fact that it dealt with *individual* particles, whereas earlier methods had always dealt with the *mean value* of the charges on a *large number* of particles.

[2] A collected report of Millikan's researches is to be found in his book "*The Electron, its Isolation and Measurement*" (Chicago, 1917). His first paper was published in August, 1909 (in the reports of the British Association Meeting at Winnipeg); Ehrenhaft's first communication appeared in March, 1909 (in the *Anzeiger der Wiener Akademie*).

gravity by $g$, and the electric strength of field (*i.e.*, the quotient of the potential difference and the distance between the plates) by $E$, we must have the following relation :—

$$\frac{v_2}{v_1} = \frac{Q\,E - m\,g}{m\,g}. \tag{1}$$

The velocity is easily determined by measuring the time taken by the particle in traversing the interval between two cross-threads in the observing microscope. Table I gives the values in seconds found in one experiment by Millikan for the times ($t_1$) required for the downward motion, and for the times ($t_2$) required for the upward motion between the cross-threads ; in this case the distance between the threads corresponded to an actual interval of rise or fall of 0·5222 cm.

TABLE I

| $t_1$ | $t_2$ | $t_1$ | $t_2$ |
|---|---|---|---|
| 13·6 | 12·5 | 13·8 | 34·6 |
| 13·8 | | 13·7 | |
| 13·4 | 12·4 | 13·8 | 34·8 |
| 13·4 | 21·8 | 13·6 | 16·0 |
| 13·6 | 34·8 | 13·5 | 34·8 |
| 13·6 | | 13·4 | 34·6 |
| 13·7 | 84·5 | 13·8 | 21·9 |
| 13·5 | 85·5 | 13·4 | |
| 13·5 | | Mean value 13·595 | |

As the table shows, the velocity of the downward motion of the particle is always the same. On the other hand, with a given potential difference, the values of $t_2$ (and hence of $v_2$ also) vary intermittently, whilst earlier values recur later on during the up-and-down motion.

The jumps in the values of $t_2$ must obviously be due to *collisions* between the positively charged oil droplet and *air ions* resulting in an alteration in the charge of the moving particle. When a negative ion is taken up, the upward motion is retarded and hence the value of $t_2$ increased, whereas a decrease in the value of $t_2$ is due to the addition of a positive ion or the loss of a negative ion.

Now denoting by $Q'$ the charge on the ion, and by $v_2'$ the velocity of the upward motion after addition of this ion, we have in accordance with eqn. (1)

$$\frac{v_2'}{v_1} = \frac{Q\,E + Q'\,E - m\,g}{m\,g}. \tag{2}$$

Hence, on subtracting eqn. (1) from eqn. (2) we find

$$\frac{v_2' - v_2}{v_1} = \frac{Q'\,E}{m\,g}. \tag{3}$$

In Table II the first column contains the values found for the velocity of the upward motion by dividing the interval of rise or fall, 0·5222 cm., by the values of $t_2$ given in Table I.[3] In the second column we find, in the first place, the differences between the successive numbers in the first column.

TABLE II

| Velocity of the upward motion. | Difference. |
|---|---|
| $\frac{0{\cdot}5222}{12{\cdot}45} = 0{\cdot}04196$ | |
| | $\frac{0{\cdot}01806}{2} = 0{\cdot}00903$ |
| $\frac{0{\cdot}5222}{21{\cdot}8} = 0{\cdot}02390$ | |
| | $\frac{0{\cdot}00885}{1} = 0{\cdot}00885$ |
| $\frac{0{\cdot}5222}{34{\cdot}7} = 0{\cdot}01505$ | |
| | $\frac{0{\cdot}00891}{1} = 0{\cdot}00891$ |
| $\frac{0{\cdot}5222}{85{\cdot}0} = 0{\cdot}00614$ | |
| | $\frac{0{\cdot}00891}{1} = 0{\cdot}00891$ |
| $\frac{0{\cdot}5222}{34{\cdot}7} = 0{\cdot}01505$ | |
| | $\frac{0{\cdot}01759}{2} = 0{\cdot}00880$ |
| $\frac{0{\cdot}5222}{16{\cdot}0} = 0{\cdot}03264$ | |
| | $\frac{0{\cdot}01759}{2} = 0{\cdot}00880$ |
| $\frac{0{\cdot}5222}{34{\cdot}7} = 0{\cdot}01505$ | |
| | $\frac{0{\cdot}00891}{1} = 0{\cdot}00891$ |
| $\frac{0{\cdot}5222}{21{\cdot}85} = 0{\cdot}02396$ | |

Table II shows us, however, that the above differences always have one of two values, of which the one is exactly twice the other (in other experiments we find also triple and quadruple values). In the case of a given droplet the

[3] When two approximately equal values of $t_2$ follow successively in Table I, the mean value has been taken for Table II.

quantities $m$, $v_1$, $E$, and $g$ are to be regarded as constants and hence it follows, in accordance with eqn. (3), that the *charges on the air ions* must always represent *integral multiples of an elementary charge.* Thus the data of Table II afford a *direct experimental proof of the atomistic structure of electricity.*

Moreover, the actual charge on the droplet ($Q$) may readily be compared with the ionic charge ($Q'$), and hence with the elementary quantum. We have only to start from eqn. (1), which gives the relation

$$\frac{v_2 + v_1}{v_1} = \frac{Q\,E}{m\,g}; \tag{4}$$

whence, taking eqn. (3) into account, we find

$$\frac{v_2 + v_1}{v_2' - v_2} = \frac{Q}{Q'}. \tag{5}$$

Since the velocities $v_1$, $v_2$, and $v_2'$ can be measured, the charge $Q$ can therefore be compared with the smallest charge, as integral multiples of which the ionic charges manifest themselves. Millikan found that the *droplet's charge* also always proves to be an *exact integral multiple of the same elementary charge.* With the greatest accuracy he was able to detect all integral multiples of the elementary charge up to about 50, and he showed that the number of elementary quanta associated with an oil droplet could even amount to 150. As the multiplying factor increases, the accuracy of the count of course decreases, for the limits of error in this method are about a half *per cent.*, so that, with a charge of more than 200 elementary quanta, the error is already more than one electrical quantum.

Now it is evident from eqn. (1) that the absolute magnitude of the *elementary quantum of electricity* itself can be calculated, as soon as the *weight of a droplet* is known. Moreover, this weight can be determined with the aid of a relation connecting the constant velocity of fall with the radius of the particle. As already found by *Stokes* in 1845, the velocity of fall of a sphere of radius $a$ and density $\varrho$ in a medium of density $\varrho'$, and with a coefficient of viscosity $\eta$, is given by the formula

$$v = \frac{2}{9}\frac{g\,a^2\,(\varrho - \varrho')}{\eta}. \tag{6}$$

Millikan and his collaborators now arranged the experiments that served to determine the weight of a droplet in such a way that the assumptions, upon which the validity

of Stokes's law is based, were satisfied with the greatest possible exactness.[4] In so far, however, as these assumptions could not be completely satisfied, the errors thereby caused were again accurately investigated. By means of the law of fall thus modified,[5] the radius of the droplets could be calculated with the greatest accuracy, and therefore, the density being known, their exact weight also followed.

Measurements extending over many years enabled Millikan in this way to determine the value of the *elementary quantum of electricity* (now generally denoted by $e$) with an *accuracy of one in a thousand ;* as a mean of his experimental results he found

(7) $e = 4{\cdot}774 \times 10^{-10}$ electrostatic units ($\pm$ $0{\cdot}005 \times 10^{-10}$).

It does not here seem necessary to enter any further into the details of Millikan's measurements, for in later sections various other methods will be described, which, quite independently of his researches, lead indirectly to exactly the same value for $e$.[6]

Historically, it may be noted that the first direct determinations of the elementary quantum of electricity were made in 1897 and 1898 by *Townsend* and *J. J. Thomson*. These researches, which gave for $e$ an approximate value of

[4] It was particularly important that the oil droplets should be perfectly spherical—a result obtained by special artifices.

[5] For further details, *vide* Millikan's book quoted above.

[6] It also follows from this that objections which have been raised against Millikan's method can in no way call in question the existence of an elementary quantum of electricity, which has been definitely proved by numerous other methods.—Objections have been made to Millikan's work by *Ehrenhaft* and his pupils, who have experimented with spherical particles of various substances (mercury, silver, gold, oil, *etc.*) in various chemically inert gases (nitrogen, argon) ; the radii of the particles lay between $5 \times 10^{-5}$ cm. and about $5 \times 10^{-6}$ cm. (the limit of observability). From these researches, some of which were published before Millikan's first paper, Ehrenhaft concluded that as the radius of the particle (and therefore its electrostatic capacity) decreased, the charge also decreased to such an extent that it might be considerably less than the elementary quantum. Subsequently Ehrenhaft discovered that, when the particles had a *radioactive* content, the velocity of their vertical movements varied *continuously*, despite the constancy of the electrical field. This he regards as a further disproof of the assumption of a universally discontinuous variation in charge of the order of magnitude of an elementary quantum.—On the other hand, Ehrenhaft's results have been ascribed to conceivable *variations in density*, which might be due to a spongy or flaky structure of extremely minute particles (smaller than Millikan's oil droplets) or to an adsorption of gas. *Cf.* Ehrenhaft's comprehensive papers in the *Ann. der Phys.*, **56** (1918), p. 1, and **63** (1920), p. 773, and in the *Zeitschr. f. Phys.*, **45** (1927), p. 557 ; also Th. Sexl, *Zeitschr. f. Phys.*, **16** (1923), p. 34. For the opposing point of view, *cf.* the summaries in *Die Naturwissenschaften* by R. Bär [**10** (1922), pp. 322 and 344], and by E. Regener [**11** (1923), p. 17].

$3 \times 10^{-10}$ electrostatic units, were not concerned, however, with individual particles of matter, but with slowly falling *clouds of mist* in which *air ions* formed condensation nuclei for the drops of water. With the aid of Stokes's law of fall, the radius of the water drops could be derived from the velocity of fall of the mist, and hence, for a known weight of mist, the number of drops could be ascertained. Since, furthermore, the total electric charge on the cloud could be measured, the charge on an individual ion could be calculated.[7] By arbitrarily identifying this charge with an elementary quantum, a value of the right order of magnitude was obtained for the latter.

### § 87. Electrolysis, and the Mass of the Hydrogen Atom.

Long before the absolute values of atomic masses were known to physicists, chemists had been in a position to determine the *relative weights of the atoms*, *i.e.*, the *ratios* existing between their masses, with the help of the so-called *Law of Multiple Proportions*, which was established by *Dalton* in 1805. As Dalton discovered, and as he formulated in this well-known law, a definite *number*, characteristic of the element, can be allotted to every chemical element in such a way that the quantities of the elements contained in a chemical compound are related to each other as *integral multiples* of the numbers characteristic of the elements concerned.

If, following convention, we make the characteristic number for *oxygen* equal to 16 *exactly*, we obtain then the so-called *atomic weights of the elements*. The smallest atomic weight is that of hydrogen—1·008 (or, more accurately, 1·0077)—the largest that of uranium (238·2). We need not enter here further into the methods by which chemists determine atomic weights; we shall have something to say about these methods later (in connection with Avogadro's Law).

The number by which the figures representing the atomic weights must be divided, in order to obtain the absolute masses of the atoms in grams, is nowadays generally known as the *Loschmidt number*. It is so called in honour of the physicist who, in 1865, was the first to determine approximately the absolute magnitude of atoms, by means of a theoretical method based on the gas laws (which we shall discuss later). Hence, denoting the mass of the hydrogen

[7] For further details, *vide*, *e.g.*, Millikan's book, ch. iii.

atom in grams by $m_{\text{H}}$ and the Loschmidt number by $L$, we have

$$m_{\text{H}} = \frac{1{\cdot}0077}{L} \text{ grams.} \tag{1}$$

In like manner, we obtain the mass of any given atom by dividing the atomic weight of the element in question by the Loschmidt number. The number of grams of a substance equal to the number which represents the atomic or the molecular weight of the substance is known as a *gram-atom* or as a *mol* (gram-molecule) respectively. Thus, for example, a gram-atom of oxygen is 16 grams, a mol of water about 18 grams, and so on. The Loschmidt number therefore gives the number of atoms in the gram-atom, or of molecules in the mol of any given substance. If, now, we know the value of the elementary quantum of electricity, we are then able to calculate the Loschmidt number with great accuracy, on the basis of the results of *electrochemical measurements*.

In 1833 *Faraday* derived the *fundamental law of electrochemistry* from his observations on *electrolysis*. This law states that the *amount* of a particular substance *produced* electrolytically in a given time is completely determined by the *current* and directly *proportional* to it. Furthermore, the amounts of the constituents liberated from different electrolytes by a current of given strength are "*chemically equivalent*," *i.e.*, they are related to each other as the quotients of the atomic weight and the chemical valency. Thus a current which liberates 1·008 grams of univalent hydrogen in a unit of time will liberate 8 grams of bivalent oxygen of atomic weight 16 in the same interval, *etc.*

The fundamental law of electrochemistry can be very simply explained on the assumption that, in an electrolyte, the molecules are entirely or partially split up, or *dissociated*, into oppositely charged components—the so-called *ions*—and that the conduction current in the electrolyte consists in a *transport* of such ions. The positive ions thus wander in the direction of the current to where it leaves the liquid, at the so-called *cathode*, whereas the negative ions wander in the opposite direction to where the current enters the liquid, at the so-called *anode*. Faraday's Law then follows on the hypothesis that the charge on all univalent ions is one elementary quantum of electricity, that on bivalent ions two, and that on trivalent ions three elementary quanta, *etc.*

Denoting the number of atoms liberated in a unit of time at the cathode by $N$ and the valency by $z$, the current, which

is the quantity of electricity passing the cathode in unit time, is given by

(2) $$I = N z e.$$

On the other hand, the mass of a single atom is equal to the quotient of the atomic weight and the Loschmidt number, *i.e.*, to the quantity $A/L$, where $A$ is the atomic weight. The amount of the substance in question liberated in a unit of time is therefore determined by the relation

(3) $$M = \frac{N A}{L},$$

or, in view of eqn. (2), by

(4) $$M = \frac{A I}{z L e}.$$

Now this equation is an actual expression of the fundamental law of electrochemistry. For, since $L$ and $e$ are universal constants, it follows on the one hand that, with a given substance, and therefore with given values of $A$ and $z$, the mass ($M$) liberated in unit time is proportional to the current ($I$); and on the other hand it follows that, with a given current, the value of $M$ is proportional to the quotient of atomic weight and valency.

Since, in accordance with what we have said above, the quantity $M/A$ represents the number of gram-atoms, we also recognize from eqn. (4) that the *product of the Loschmidt number and the elementary quantum of electricity* simply represents the *current*, measured in electrostatic units, which *liberates one gram-atom of a univalent element* in a unit of time.

Very accurate measurements, in which principally silver has been used, have demonstrated that a gram-atom of a univalent substance would be deposited in a unit of time by a current of 96,494 ampères. In order to express this current in electrostatic measure, we have to multiply it by one-tenth of the velocity of light [according to § 64 eqn. (9)]. Since the velocity of light is equal to $2{\cdot}998 \times 10^{-10}$ cm. sec.$^{-1}$, we find thus

(5) $$L e = 2{\cdot}893 \times 10^{14} \text{ E.S.U.}$$

Hence, introducing the value already given for the elementary quantum of electricity, we obtain for the *Loschmidt number*

(6) $$L = 6{\cdot}060 \times 10^{23},$$

with an error which can scarcely be more than 1 in 1,000.

It follows from eqn. (1) that the *mass of the hydrogen atom* is

$$m_{\mathrm{H}} = 1{\cdot}662 \times 10^{-24} \text{ grams.} \tag{7}$$

For the ratio of charge to mass, *i.e.*, for the so-called *specific charge of the ionized hydrogen atom*, we have, in accordance with eqn. (1),

$$\frac{e}{m_{\mathrm{H}}} = \frac{L\,e}{1{\cdot}0077}, \tag{8}$$

or, by eqn. (5),

$$\frac{e}{m_{\mathrm{H}}} = 2{\cdot}871 \times 10^{14} \text{ E.S.U./gram.} \tag{9}$$

### § 88. Negative Radiation and the Electron.

In 1869 *Hittorf* discovered the remarkable rays which are emitted by the cathode in highly evacuated Geissler tubes, and are therefore known in general as *cathode rays*. Their properties were more closely investigated, particularly by *Crookes* (1879), who made the surprising discovery that these rays, which ordinarily travel in *straight lines*, are very easily *deflected* by a *magnet*. Moreover, *Goldstein* (1876) showed that they may also be *deflected* by an *electric field*.[1]

As was first pointed out by *J. J. Thomson* and *Kaufmann* (1897), all these properties of the cathode rays can be simply explained on the assumption that they consist of quickly moving *electrically charged particles*, which at the same time possess *inertial mass*. Their rectilinear propagation follows from their inertia, and the observed deflections both in a magnetic and in an electric field necessarily follow from their electrical charge ; the sense of the deflection in both cases shows that the *charge* must be *negative*. Unlike rays of light or heat, which are forms of electromagnetic waves, the cathode rays therefore represent a *material radiation* consisting of a stream of inertial particles.

Now (as was shown in § 71) it follows from the theory of electricity, without the aid of any kind of atomistic hypothesis, that, in a *transverse homogeneous magnetic field*, a moving electrically charged mass must describe a *circular path* with constant velocity in a plane at right angles to the

[1] The study of the cathode rays was greatly aided by an artifice of *Lenard's*, which consisted in letting the rays escape from the tube through a piece of thin metal foil inserted in the wall of the tube.

magnetic lines of force. In accordance with § 71 eqn. (11), the radius of the circle is given by the simple relation

$$r = \frac{c\ v}{\gamma H}, \tag{1}$$

where $H$ is the magnitude of the magnetic strength of field, $\gamma$ the specific charge of the moving mass, and $v$ its velocity.

It follows likewise from the theory of electricity (according to § 71), irrespective of any kind of atomistic considerations, that, in a *transverse homogeneous electric field*, a moving charged body describes a *parabola*, the curvature of which depends only on the specific charge and the velocity, apart from on the electric field-strength. If we take the original direction of the path as $x$-axis, the direction of the electric force as $z$-axis, and the position of the moving body at the time $t = 0$ as origin, we have then [by § 71 eqn. (3) ] the relation

$$\frac{z}{x^2} = \frac{1}{2}\frac{\gamma}{v^2} E, \tag{2}$$

where $E$ is the magnitude of the electric field-strength.

As a matter of fact, observations show that the particles in the cathode rays do move in a circular path in a transverse magnetic field, and in a parabola in a transverse electric field. Hence, if we apply the two equations to cathode rays, and regard $\gamma$ and $v$ as unknown, we may calculate the velocity and the specific charge of the particles from the known measured values of the magnetic and electric strengths of field.

Exceedingly large values are obtained for the *velocity*, being of the order of magnitude *of the velocity of light*. Naturally, the velocity depends on the potential applied to the cathode tube (*viz.*, on the potential difference through which the particles travel). If $V$ be the tube potential and $Q$ the charge on a cathode particle, the product $Q\ V$ must, of course, be equal to the kinetic energy of the particle as it leaves the tube, *i.e.*, to $\frac{1}{2}\ mv^2$. Hence

$$v = \sqrt{2\gamma V}. \tag{3}$$

Thus the velocity varies as the square root of the tube potential.

The value of the *specific charge*, as obtained from measurements according to eqns. (1) and (2), is given by

$$\gamma = 5{\cdot}3 \times 10^{17}\ \text{E.S.U./gram}. \tag{4}$$

On comparing this value with that obtained for the specific

charge of the ionized hydrogen atom, we see that it is about 1,800 times as large. If we make the very likely assumption that the *charge on a cathode ray particle* amounts to *one elementary quantum of electricity*, we must then ascribe to the particle a *mass* 1/1800th of that of the *hydrogen atom*.

When this assumption was first suggested, it appeared to physicists an extremely bold one, because, up to about the end of the 19th century, the mass of the hydrogen atom had been regarded as the smallest mass possible. In 1896, however, the discovery of the *Zeeman effect* (to be discussed later), which consists in the resolution of spectral lines in a magnetic field, led to the same conclusion. Moreover, the subsequent progress of theoretical physics has appeared entirely to confirm all deductions made on the hypothesis of smallest particles with a mass 1/1800th of that of the hydrogen atom. These particles with a negative charge of one elementary quantum of electricity are now generally known as *electrons*.[2] It follows from eqn. (4) and the known value of the elementary quantum of electricity that the *mass* of an electron is given by

$$m = 9{\cdot}0 \times 10^{-28} \text{ grams.} \tag{5}$$

Besides the artificially produced negative rays, physics is also acquainted with spontaneously occurring rays, likewise consisting of electrons, but generally possessing a much higher velocity: these rays form part of the radiation emitted by so-called *radioactive substances*. In 1896 *Becquerel* made the surprising discovery that *uranium ores* constantly emit rays without the intervention of any outside agency. These rays are able to blacken a photographic plate through an opaque covering, and also to ionize the air, thus rendering it conducting, so that a charged electroscope set up close at hand is discharged. Two years later (1898), *Pierre* and *Marya Curie* found that the so-called Becquerel rays (as they were originally termed) are, in point of fact, principally emitted by compounds of a metal contained in the pitchblende. At that time this metal was still unknown, but, after tedious researches, M. and Mme. Curie isolated it from pitchblende and named it *radium*.[3] A closer investigation of the "radioactive radiation" emitted by radium and other substances showed that it consists of *three* entirely different kinds of rays, which are

[2] The term "electron" is due to Stoney (1881).

[3] *Metallic* radium was not obtained until 1910, by Mme. Curie; up till then only salts of radium were known.

distinguished as *alpha-*, *beta-*, and *gamma-rays ;* a separation of the three kinds of rays is rendered possible by their *different behaviour in a magnetic field.* Whereas the $\gamma$-rays are not deflected at all, and are thus to be regarded as a radiation of electromagnetic waves, the $\alpha$- and the $\beta$-rays are found to be material rays that are deflected by a magnet. Furthermore, it appears that the $\alpha$-rays (which will be discussed later) carry positive charges, whereas the $\beta$-rays carry negative charges.

The *velocity* and *specific charge* of the *beta-rays* could be determined by the same method as in the case of cathode rays. The value obtained for the specific charge was the same as for cathode rays, which apparently proved the $\beta$-rays to consist likewise of *electrons.* They are thus essentially identical with cathode rays, but they are distinguished from them by a *larger velocity,* for which values ranging from 30 *to* 99·8% *of the velocity of light* have been obtained.

The motion of $\beta$-particles is the *most rapid material motion* known in physics, and for this enormous velocity so closely approaching that of light [4] are clearly revealed the *deviations from classical mechanics* which, according to the *theory of relativity,* are to be expected in the case of such rapid motions. It follows from this theory—the subject will be fully discussed later—that *mass* depends on *velocity* in accordance with the relation·

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}, \tag{6}$$

where $m_0$ denotes the value of the mass for velocities small compared with the velocity of light ($c$). This formula shows that, with a velocity of $v = c/2$, the mass is some 1·15 times as large as $m_0$, whilst, with a velocity of 99·8% of the velocity of light, the mass is as much as 16 times as large. Measurements by *Kaufmann* and other investigators have actually established the increase in mass with increasing velocity, as demanded by the relativity theory, and have also confirmed eqn. (6) quantitatively.[5]

Essentially the same as cathode and $\beta$-rays are the rays which are a consequence of the so-called *photo-electric effect.* In continuation of an earlier research by *Hertz,*

[4] It appears from the theory of relativity that velocities greater than the velocity of light are impossible, as is also shown by eqn. (6).

[5] An increase in mass connotes a decrease in $\gamma$.

*Hallwachs* in 1888 discovered that a *metal plate* becomes *positively charged* when illuminated with *ultra-violet light*. *Lenard* recognized the cause of this phenomenon to be a negative radiation emitted by the illuminated metal. This radiation was found to have the same specific charge as cathode rays, so that obviously it must also consist of *electrons*. Their velocity, however, proved to be much smaller than that of cathode rays, amounting only to some thousandths of the velocity of light. Of the same order of magnitude are the velocities of the *electrons* emitted by *glowing* bodies, as was first observed by *Richardson* in 1903.

Every material radiation of definite velocity and definite specific charge may be characterized by a definite value of the electric potential, which may be called the *equivalent potential* and, when multiplied by the elementary quantum of electricity, gives the kinetic energy of the radiated particle. Thus, in accordance with eqn. (3), we have the relation

$$V = \frac{v^2}{2\gamma}, \tag{7}$$

or, by § 64 eqn. (10), when the potential is expressed in volts,

$$V = 150\frac{v^2}{\gamma} \quad (V \text{ in volts, } \gamma \text{ in electrostatic measure}). \tag{8}$$

In the case of cathode rays we find, of course, as equivalent potential the tube potential, or about 30,000 volts for velocities amounting to one-third of the velocity of light. The potential corresponding to the fastest $\beta$-rays would be about 4,000,000 volts, for $v$ is then almost equal to $c$, and $\gamma$ is 16 times smaller than in the case of cathode rays (in consequence of the increase in mass).

### § 89. Positive Radiation.

The negative rays which we have hitherto been discussing find their counterpart in the so-called *canal rays* discovered by *Goldstein* in 1886. Like the cathode rays, the canal rays emanate from the cathode in a highly evacuated discharge tube, but travel in the opposite direction to the cathode rays.[1] It was, moreover, later observed that the anode itself emits positive rays, known as *anode rays*,[2] and there is no essential difference between them and canal rays. Their velocity and specific charge can be determined

[1] In order that the rays may escape from the cathode, the latter must be perforated by small canals; hence the name "canal rays."

[2] The anode rays were discovered in 1906 by *Gehrcke* and *Reichenheim*.

by the same method as in the case of negative rays, the sense of deflection clearly demonstrating that the particles in canal and anode rays are *positively charged*. Their *velocity* is found to be a few thousandths of the velocity of light.

Experiment shows the value of the *specific charge* to be much smaller than in the case of the negative rays ; furthermore, this value depends on the particular gas of which traces remain in the evacuated tube. The largest value of the specific charge is obtained when the tube is filled with hydrogen before evacuation ; in that case the value is identical with that obtained from electrolytic measurements for the specific charge of the hydrogen ion. If the tube be filled with mercury vapour before evacuation, the specific charge is found to have a value 1/200th of the above, corresponding to the fact that the atomic weight of mercury is 200. The particles in positive rays must therefore be regarded as *electrically charged atoms*, the charge on each amounting to one, or sometimes to several elementary quanta.[3]

Now, in transverse electric and magnetic fields the particles in a beam of material rays must be *deflected*, and, if the electric field be arranged at right angles to the magnetic field, it can be shown that all particles with the same specific charge will strike a plate, placed at right angles to the beam, along a *parabolic curve*. Hence, if a positive radiation be composed of atoms with different specific charges, these must form parabolas of different curvatures, and upon this fact rests the possibility of *positive ray analysis*, as was first recognized and applied by *J. J. Thomson*.

The method of positive ray analysis has been developed by *Aston* (1919) into one of the most important methods of experimental research.[4] Owing to the focussing effect obtained, Aston was able to produce much greater resolution with his form of apparatus than had been possible with the original method. In this way the masses of arbitrary particles can be accurately compared with one another ; and in addition to some other very important experimental

[3] In addition to charged atoms, we may also meet with charged atomic complexes, *e.g.*, molecules of hydrochloric acid, which possess a positive charge of one elementary quantum, and so on.

[4] In Aston's method the electric and magnetic fields are applied so that the deflections produced are exactly opposed to each other. A very thin ribbon of rays is spread out by the electric field into an "electric spectrum," and a group of the deflected particles is allowed to pass through the magnetic field. It may be shown that all particles with the same mass strike a suitably placed photographic plate along a sharply defined straight line.

results which we shall discuss fully later, Aston's researches established the most important fact that the *masses of arbitrary positive ray particles* are *always approximately integral multiples of the mass of the hydrogen atom* (or, more accurately speaking, almost always whole numbers in terms of the mass of the oxygen atom, taken as 16).

Like cathode rays, canal and anode rays are artificial products of experimental physics. A *natural* counterpart to the negative $\beta$-rays is represented by the *positive alpha-rays* which, as we have already mentioned, are emitted similarly to the $\beta$-rays by radioactive substances. Their positive electrical nature is proved by the fact that they are deflected by a magnet, but in the opposite sense to $\beta$-rays.

The *velocity* of $\alpha$-rays generally lies between *five* and *seven per cent. of the velocity of light*, varying with the chemical nature of the radioactive substance emitting them. Measurements of the *specific charge* yield a value of

$$\gamma = 1{\cdot}44 \times 10^{14} \text{ E.S.U./gram.} \tag{1}$$

Thus the specific charge of the $\alpha$-particle is *half as large as that of the ionized hydrogen atom*. Now it is a fact of great importance that the *$\alpha$-particles* emitted by a radioactive substance can be *counted*, for the *charge on an individual $\alpha$-particle* can then be directly determined. This is rendered possible by a remarkable phenomenon known as *scintillation*. If we bring a radioactive preparation into the neighbourhood of a fluorescent screen (of zinc blende or a very thin section of diamond), we observe a continual flashing of discrete points of light. This suggests that each flash of light is due to the impact of a *single* $\alpha$-particle.[5] We can place the radioactive preparation behind a screen of some substance which absorbs the $\alpha$-radiation, and in which there is a small hole allowing an exactly calculable fraction of the total $\alpha$-radiation emitted by the preparation to pass through. This fraction is allowed to strike a small portion of the scintillation screen, which is observed under a microscope. In this way we can readily count the number of scintillations occurring in a definite time, and, knowing the size of the hole, we can calculate how many $\alpha$-particles are emitted every second by a given preparation. Determinations of this kind were first made by *Rutherford* and *Geiger* in 1908.

[5] This assumption is justified by the fact that the scintillation method gives the same value for the disintegration constant of radium as do other methods (which we shall discuss subsequently).

Furthermore, by cutting out the $\beta$-rays emitted simultaneously with the $\alpha$-rays by means of a magnetic field, a metal plate can be charged positively by the radioactive radiation. In order to find the electric charge on a single $\alpha$-particle, we have then only to divide the total charge conveyed to the plate in a given time by the known number of $\alpha$-particles which strike the plate in that time. Rutherford and Geiger's measurements showed that the value of the charge on an $\alpha$-particle is exactly *double* that of the *elementary quantum of electricity*. Since, on the other hand, the specific charge is half that of the ionized hydrogen atom, we must ascribe to the *$\alpha$-particle* a *mass equal to that of four hydrogen atoms*, and a *positive charge of two elementary quanta*.

On *passing through ordinary substances*, *$\alpha$-rays* are far more strongly *absorbed* than $\beta$-rays, and, moreover, the nature of the absorption is entirely different. For, whereas $\beta$-rays are absorbed by matter in accordance with an *exponential law*, the action of $\alpha$-rays (such as the blackening of a photographic plate or the ionization of air) ceases quite suddenly, as soon as the rays have penetrated a certain distance into a given medium. This distance is known as the *range*, and it depends on the chemical character of the radiating substance as well as upon the nature of the absorbing medium. The greater the initial velocity of the $\alpha$-rays on their emission from the radioactive substance, the larger is the range, which, as *Geiger* has empirically discovered, is *proportional to the cube of the initial velocity*. In ordinary air the ranges of known $\alpha$-rays lie between $2 \cdot 7$ and $11 \cdot 3$ cm.; in hydrogen the fastest $\alpha$-rays have a range of 24 cm. (at room temperature and ordinary pressure). An aluminium plate $0 \cdot 06$ mm. thick suffices to absorb completely all but the fastest $\alpha$-rays, whereas $\beta$-rays can penetrate through aluminium plates some millimetres in thickness.

The passage of $\alpha$-rays through matter has been investigated with particular success by *C. T. R. Wilson* (1912). His experiments were based upon the well-known fact that, in *supersaturated water vapour*, electrically charged particles serve as *condensation nuclei* for *water drops*. Thus, on bringing a radioactive preparation into a chamber containing supersaturated water vapour, Wilson was able to *photograph* the collection of water drops created by the passage of an $\alpha$-particle. A photograph of the streaks of mist thus produced gives also a picture of the *actual paths of the $\alpha$-particles*. In general the photographs (Figs. 59 and 60) show *straight lines*, many of which, however, appear *bent at the end* through

fairly large angles. This fact, which we shall discuss in detail later, is of fundamental importance for modern atomic theory.

Besides canal, anode, and $\alpha$-rays, a fourth kind of positive radiation is known in physics, *viz.*, the so-called *H-rays*, discovered by *Marsden* in 1914. When he passed the $\alpha$-particles emitted by radium C through *hydrogen*, in which their range is 24 cm., Marsden was able to detect scintillations on a screen even at a distance of 80 cm. These scin-



Fig. 59.



Fig. 60.

Paths of $\alpha$-rays.

tillations could obviously not be due to the $\alpha$-rays, but must have been caused by other rays which, in their turn, were produced by the $\alpha$-particles. A determination of the specific charge gave the same value for these newly discovered rays as is obtained from electrolytic measurements for hydrogen ions. From the chemical symbol for hydrogen, these rays are therefore called H-rays.

*Equivalent potentials* can easily be calculated for positive radiation. From eqn. (1) and § 88 eqn. (8) we find for $\alpha$-rays,

the velocity of which, as already mentioned, amounts to 5 to 7% of the velocity of light, a value of about two to four million volts. If we wish to express in ergs the kinetic energy of a rapid $\alpha$-particle possessing a velocity of $2 \times 10^9$ cm./sec., we have to multiply half its mass (*i.e.*, twice the mass of a hydrogen atom) by the square of its velocity. We find thus

$$E_{\text{kin}} = 1{\cdot}3 \times 10^{-5} \text{ ergs.} \tag{2}$$

The velocity of the fastest $\beta$-particles is, it is true, some 15 times as large, and therefore the square of the velocity about 200 times as large. On the other hand, their mass is some 7,000 times as small, and even when we take into account the 16-fold increase in mass for the enormous velocity of 99·8% of the velocity of light, the mass is still some hundreds of times as small as that of an $\alpha$-particle. The kinetic energy of $\beta$-particles is therefore considerably less than that of $\alpha$-particles.

### § 90. The Elementary Quantum of Action.

During an investigation of the *photo-electric effect* in 1902, *Lenard* had already observed the remarkable fact that the *velocity* of the electrons liberated by ultra-violet light is entirely *independent of the intensity of the light used.* The velocity cannot be increased by raising the intensity, but only by raising the frequency of the light. Increasing the intensity augments only the *number* of liberated electrons, but not their velocity.

Quantitative measurements of the photo-electric effect are based on the fact that the emission of the electrons from the metal plate subjected to the radiation can be prevented by giving the plate a *positive charge.* With a given frequency of the incident light, we can thus determine for a plate of a given metal a quite definite *counterpotential*, on reaching which the emission of electrons just comes to a complete stop. If $V$ be the counterpotential, observation shows that *$V$ increases according to an exact linear relation with the frequency of the light.* We may write

$$V = C\nu - P, \tag{1}$$

where $C$ and $P$ are positive constants. We find from experiment that only the constant $P$ depends on the nature of the metal,[1] whereas the constant $C$ is independent thereof and therefore represents a *universal constant.* Now the product

[1] It is true that the presence of adsorbed or occluded gases has a great influence.

$e\ V$ denotes the kinetic energy of the most rapid electrons liberated, and this suggests a multiplication of eqn. (1) by $e$. If we write $h$ for the product $C\ e$, eqn. (1) becomes

$$e\ (V + P) = h\ \nu. \tag{2}$$

Since the product $e\ V$ represents work or energy, whilst frequency has the reciprocal dimensions of time, the dimensions of the constant $h$ must be those of a *product of energy and time*. Such a product in mechanics is termed *action*, and the constant $h$ is therefore called the *elementary quantum of action*, forming thus a *fourth* elementary quantum, in addition to the elementary quantum of electricity and the masses of the electron and the hydrogen atom. Its value is found from measurements of the photo-electric effect to be

$$h = 6{\cdot}55 \times 10^{-27}\ \text{erg. sec.} \tag{3}$$

The hypothesis of the elementary quantum of action was introduced into theoretical physics by *Planck* in 1900. By means of this hypothesis Planck succeeded in solving a difficult fundamental problem in the theory of thermal radiation, which we shall discuss at length later. For this purpose he assumed that the emission of wave radiation occurs *discontinuously* in such a way that *elements of energy* of magnitude $h\ \nu$ play a part in the process.

### § 91. Light Corpuscles and Matter Waves.

In 1905 *Einstein* extended Planck's hypothesis of the discontinuous emission of radiation by introducing the concept of *light-quanta*, the energy of which is to be determined by the product of the frequency and the elementary quantum of action. Einstein assumed that every mutual *transformation between electromagnetic wave radiation and material radiation* proceeds in such a way that either a *single light-quantum* is utilized in the liberation of a *single electron*, or, conversely, a single light-quantum is produced by the kinetic energy of a single electron.

Actually it follows necessarily from Einstein's assumption [and in agreement with § 90 eqn. (2)] that, in the *photo-electric effect*, a light-quantum of the incident light must be equal to the kinetic energy with which a liberated electron leaves the irradiated metal, increased by the work performed by the electron in its passage from the metal. Einstein's hypothesis is also confirmed by the converse phenomenon to the photo-electric effect, *viz.*, the transformation of material into electromagnetic wave radiation; we

shall discuss this in a subsequent section dealing with the production of X-rays by cathode rays. Moreover, the concept of light-quanta provides a simple explanation for a principle governing the transformation of primary into *secondary wave radiation*. For if *each secondary* light-quantum be due to a separate primary light-quantum, the principle of energy teaches that the secondary frequency can never be larger than the primary; it must be *smaller*, if energy is used up for other purposes during the transformation. Secondary wave radiation is termed *fluorescent light*, and, in fact, in 1852 *Stokes* established the *rule*, known by his name, that the *wave-length* of the fluorescent light is *longer* than that of the exciting light.

In 1923 *Arthur H. Compton* made a discovery which clearly indicated that, just as material bodies, *light-quanta* must also possess *mechanical momentum*.[1] Compton found that X-rays, when scattered by loosely bound electrons, experience a decrease in frequency. The principles of the *Compton effect*, which will be discussed more fully in a later section, may, however, be simply explained on the assumption that, in the *collisions between light-quanta and electrons*, the Laws of the *Conservation of Energy* and of the *Conservation of Momentum* are obeyed, provided that we associate with each light-quantum, besides an energy $h\nu$, also a *momentum* in the direction of motion given by

$$G = \frac{h\nu}{c}. \tag{1}$$

This value for the momentum follows from the principle of the *Inertial Mass of Energy*, established by the theory of relativity. According to this principle, an energy $E$ possesses a mass $E/c^2$, so that the mass of a light-quantum must be $h\nu/c^2$. On multiplying this quantity by the velocity of propagation of light-quanta, *i.e.*, by $c$, we obtain for the momentum (defined as the product of mass and velocity) the value given in eqn. (1).[2] Furthermore, since the velocity of light is equal to the product of frequency and wave-length ($\lambda$), we may put

$$G = \frac{h}{\lambda}. \tag{2}$$

[1] This conclusion had already been drawn by Einstein in 1917 from considerations on the interaction between radiation and matter [*Physikal. Zeitschr.*, **18** (1917), p. 121].

[2] According to § 88 eqn. (6) the "stationary mass" of a light-quantum would be equal to zero. The difficulty of such a conception can best be avoided by considering the momentum of a light-quantum rather than its mass.

Again, the long well-known *Doppler effect* can readily be explained upon the assumption of corpuscular light-quanta. If a body of mass $m$ moving with a velocity $v$, emit a light-quantum *in the direction of motion*, it follows from the Law of the *Conservation of Momentum* that

$$m\,v = \frac{h\,\nu}{c} + m\,(v - \Delta v), \tag{3}$$

or,

$$m\,\Delta v = \frac{h\,\nu}{c}. \tag{4}$$

On the other hand, we obtain from the Law of the *Conservation of Energy* the relation

$$\frac{m\,v^2}{2} + E = \frac{m}{2}\,(v - \Delta v)^2 + h\,\nu, \tag{5}$$

where $E$ denotes the change in energy consequent upon the emission of the light-quantum: it is equal to $h\,\nu_0$, if $\nu_0$ be the frequency when the body is stationary. If $\Delta v$ be small in comparison with $v$, eqn. (5) thus becomes

$$h\,(\nu - \nu_0) = m\,v\,\Delta v, \tag{6}$$

or, in accordance with eqn. (4),

$$\nu - \nu_0 = \nu\,\frac{v}{c}. \tag{7}$$

Hence we find, by way of approximation, the well-known Doppler formula for the *displacement of the spectral lines*:

$$\Delta\nu = \nu_0\,\frac{v}{c}. \tag{8}$$

When the direction of emission is *opposed* to that of motion, the increase in frequency is replaced by a decrease of equal magnitude. Where an arbitrary angle is included between the two directions, we have to understand by $v$ the *component* of the velocity in the direction of motion.

The progress of physics has thus led to an evolution of the hypothetical light-quanta into *material corpuscles of light*. Indeed observations, which had already been suggested by Einstein's general theory of relativity, showed that not only is mechanical momentum to be associated with light-quanta, but also that, in a gravitational field, a *potential gravitational energy* is to be ascribed to them. As is well known, a body of mass $m$ situated on the surface of a star (or planet) of mass $M$ and radius $R$ possesses a poten-

tial energy $f\,M\,m/R$, where $f\,(=6{\cdot}68\times10^{-8})$ denotes the universal constant of gravitation. This potential energy is a measure of the work necessary to remove the body from the surface of the star completely out of the latter's gravitational field.

If, now, potential energy is associated with the light-quanta in a gravitational field, we must assume that a light-quantum radiated from the surface of a star will perform *work* against the attraction of the star of an amount given by

$$A=\frac{h\,\nu}{c^2}\,\frac{f\,M}{R}\,. \tag{9}$$

It follows from the concept of *separate* light-quanta that this work must be done by each *individual* light-quantum, and, since the energy of a light-quantum is equal to its frequency multiplied by the universal factor $h$, removal from the gravitational field must result in a *decrease in the frequency of the light* according to the relation

$$h\,\nu'=h\,\nu-A\,. \tag{10}$$

Hence, denoting the alteration in frequency by $\Delta\nu$, we have by eqn. (9)

$$\frac{\Delta\nu}{\nu}=-\frac{f\,M}{R\,c^2}. \tag{11}$$

This formula describes the so-called *displacement towards the red*,[3] which should be observable for *spectral lines* originating from *stars* relatively to the corresponding lines with a *terrestrial* origin. Eqn. (11) was first derived in 1915 by *Einstein* on the basis of his *general theory of relativity*, quite independently of the hypothesis of light-quanta. For the earth the ratio $f\,M/R\,c^2$ is about $6\times10^{-10}$, and for the *sun* approximately $2\times10^{-6}$. Thus, in the case of the sun, the effect first predicted by Einstein lies on the borders of observability ; it was, however, apparently confirmed by observations made in 1924 by *Adams* [4] on the *companion of Sirius*, where the ratio is about 30 times as large as for the sun. Moreover, another consequence of the hypothesis of ponderable light corpuscles has been experimentally confirmed,

[3] We speak of a displacement towards the red because, within the visible portion of the spectrum, the frequency decreases from the violet towards the red.

[4] W. S. Adams, *Proc. Nat. Acad. U.S.A.*, 11 (1925), p. 382.

*viz.*, that a *ray of light*, on passing close by the sun, is *curved* by the gravitational action of the latter.[5]

The hypothesis of light-quanta (or, as they have recently been named, of *photons* [6]) ascribed to light the essential properties of matter, such as discontinuous structure, momentum, and weight. This suggested the converse process, the associating with *matter* of the chief characteristic of light—its *wave-like nature*. In 1924 the corpuscular conception of light was in this way completed by the birth of *Louis de Broglie's wave theory of matter*.

According to this theory, into which we shall enter more fully later, every *moving* particle of matter is to be associated with a *material wave* and especially with a characteristic *wave-length*. This wave-length shall bear the same relation to the momentum of the particle as does the wave-length of a light-quantum to its momentum ; *i.e.*, in accordance with eqn. (2), the relation

$$\lambda = \frac{h}{m\,v} \tag{12}$$

shall be satisfied also for material waves, where $m$ is the mass of the particle and $v$ its velocity.

For an *electron*, the value of the quotient $h/m$ is 7·27. In the case of electrons with a velocity of the order of $10^7$ to $10^9$ cm./sec., the wave-lengths of the corresponding material waves thus agree with those of *X-rays*.

In point of fact, *Davisson* and *Germer* in 1927 were able to show that, just as X-rays, the electrons emitted from glowing wires are *selectively reflected by crystals*, *i.e.*, they are reflected in certain definite directions determined by the *structure* of the crystals. These and similar experiments, which we shall discuss in a later section (§ 121), prove conclusively that the *corpuscular* and *wave theories* are no longer to be regarded as the foundation stones of different branches of physical science, but that they represent *complementary* characteristics of one general and uniform portrayal of natural processes.

## § 92. The Nuclear Atom.

As we have seen, the phenomena of electrical radiation render manifest the existence, on the one hand, of particles

[5] It is true that, according to the general theory of relativity, the total curvature of a ray of light comes out at twice as large as on the basis of ponderable light-quanta, unless some further special assumptions are made. Observations hitherto effected depend on photographs of the immediate vicinity of the sun during a total solar eclipse.

[6] The name of "photon" is due to G. N. Lewis, *Nature*, **118** (1926), p. 874.

with a *negative charge* of one elementary quantum of electricity and with a mass invariably equal to 1/1,847th of that of the hydrogen atom, and, on the other hand, of corpuscles with a *positive charge* of one or of a few elementary quanta and possessing a *mass* which is always approximately an *integral multiple* of that of the hydrogen atom.

It would appear that *atoms* must be *composed* of these negative and positive corpuscles, and important clues as to the nature of this structure are afforded by the previously described observations on the *passage of alpha-rays through matter*. These observations show that an $\alpha$-particle can describe *rectilinear paths* of a few centimetres in air. A cubic centimetre of air, however, contains approximately $5 \times 10^{19}$ atoms, as follows from the known density and from the known value of the Loschmidt number.[1] On the other hand, it has long been known from the kinetic theory of gases that the diameter of an atom must be of the order of magnitude of $10^{-8}$ cm., so that, in air, about 1/100,000th part of the total volume is occupied by the atoms, and a line 1 cm. long therefore passes through thousands of atoms. Hence an $\alpha$-particle is capable of *penetrating through many thousands of atoms* without its direction being appreciably affected. Occasionally, however, as Wilson's photographs show, *a single atom deflects the $\alpha$-particle through a very large angle.*

This remarkable fact can be readily explained, on the basis of *Coulomb's Law*, by the assumption that the deflection is caused by an *electrical force* exerted on the electrically positive $\alpha$-particle by an *atomic nucleus*, the *volume* of which constitutes *only a very small part of the total volume of the atom.* We might, at first, equally well suppose the electrical force to be one of attraction as of repulsion, and hence the atomic nucleus to be either negative or positive.

Now it follows from the principle of action and reaction that the extraordinarily large deflections of the $\alpha$-rays actually observed are only possible if the *mass of the nucleus* is considerably *larger* than the mass of the deflected $\alpha$-particle. Observations show, in fact, that the magnitude of these large deflections increases with the atomic weight of the deflecting substance. We must therefore assume that the *mass of the atom* is principally *concentrated in the nucleus.*

Again, since the mass of the electrons is vanishingly small in comparison with that of the positive electrical particles, we must suppose that the nuclei of the atoms are electrically

[1] The number of molecules is half as large.

*positive.* Inasmuch as, in general, the atoms as a whole are electrically *neutral*, this positive charge must be *compensated* by negative charges; and, because negative particles are found *only* as electrons in cathode and $\beta$-rays, the positive nuclear charge must obviously be compensated by *electrons* surrounding the nucleus. Provided that the *motion* of these electrons *around the nucleus* be sufficiently rapid, the electrical force of attraction which would otherwise cause them to unite with the nucleus will be counterbalanced by the *centrifugal force.* We obtain in this way a *model of the atom* which was first introduced into theoretical physics by *Rutherford* in 1911. According to this model, the atoms consist of positively charged nuclei around which, to borrow a frequently used simile, electrons revolve like planets around the sun.[2]

The *simplest* imaginable atomic model is obviously one in which a *single electron* revolves about an *atomic nucleus* charged with *one positive elementary quantum.* The idea at once suggests itself to associate this model with the *lightest* of all atoms, the *hydrogen atom*, and it is actually found that all conclusions which may be theoretically derived from this assumption are completely confirmed by experiment. The positive fundamental particles of matter, which are now known generally as *protons*, thus prove to be identical with *hydrogen nuclei.*

### § 93. Characteristic Potentials and Spectral Terms.

If atoms are composed of particles between which forces of attraction and repulsion depending on distance are at work, it may well be assumed that atoms possess an *internal energy*, and that this may be increased by the addition of energy from without or, in some circumstances, decreased by the loss of energy to external bodies. That *atoms* can actually *take up energy* appears clearly from experiments on

[2] Rutherford based his atomic model, which he had developed a year before Wilson's researches, on his own observations on the passage of $\alpha$-rays through metal foils.—To-day only a historical significance is attached to *J. J. Thomson's model of the atom* which, at that time, was a subject of much discussion. Thomson supposed the atom to consist of a uniformly dense spherical volume charge of positive electricity, within which electrons described circular orbits about the centre of the sphere. The force acting on an electron is then directly proportional to its distance from the centre (because only a portion of the spherical charge is effective). Although the assumption of such forces proved most fruitful for the explanation of dispersion and magneto-optical phenomena, Thomson's atomic model had to be relinquished when observations on the passage of $\alpha$-rays through matter rendered it untenable.

*collisions between atoms and free electrons.* At the same time, however, these experiments reveal the extremely important fact that an atom can take up, not arbitrary amounts, but only *quite definite quantities of energy.* In particular it is found that, provided that no abnormal conditions are present and the atom is thus in its normal state, an atom is unable to absorb any amount of energy less than a certain *minimum.* This minimum is known as the *excitation energy of the atom.*

The phenomena of *electron impacts*, which were first investigated by *Franck* and *Hertz* in 1913, must for two reasons obviously afford much simpler and clearer experimental results than collisions between atoms. On the one hand, the kinetic energy of a single electron can be exactly determined from a measurement of the electrostatic potential. On the other hand, the mass of an electron is negligible as compared with that of an atom, and hence the exactly measurable kinetic energy lost by an electron can at once be equated to the increase in internal energy experienced by the atom involved in the collision.

Now, as long as the kinetic energy of an electron is smaller than the energy required to excite an atom, the kinetic energies of the electrons will not be diminished by the collisions. In this case, *only "elastic" impacts* are possible. *Inelastic impacts* (in which the colliding electron loses energy) will first be possible when a critical value of the potential is reached, such that, when multiplied by the elementary quantum of electricity, the product is equal to the excitation energy ; this critical value is known as the *excitation potential.*

For many elements, it has been possible to make direct experimental determinations of the excitation potential. Using suitable apparatus, the electrons are made to pass through the gas or vapour, and the resulting current is measured with a galvanometer, while the applied potential difference is also observed. If the current be plotted on a graph against the potential, the curve exhibits a strongly marked discontinuity for the potential at which inelastic impacts first occur. In this way the excitation potentials for several elements have been found to have values of some few volts : *e.g.*, for lead the value is 1·26, for sodium 2·12, for mercury 4·9 volts, *etc.* The excitation energy in *ergs* is obtained on multiplying the excitation potential, expressed in absolute units (volts divided by 300), by the elementary quantum of electricity ($4{\cdot}77 \times 10^{-10}$). The

relation between excitation energy and potential is thus given by

(1) $$E = 1{\cdot}591 \times 10^{-12} \text{ ergs per volt,}$$

so that for sodium, for example, the excitation energy amounts to $3{\cdot}37 \times 10^{-12}$ ergs.

The potential at which the electron impacts first produce ionization of the atoms of the gas can also be fairly sharply determined for many elements. This potential is called the *ionization potential;* it corresponds to the so-called *ionization energy*, or the work necessary for the *removal* from the atom of the *most loosely bound* of its *electrons*. The ionization potentials of the elements prove to be about two or three times as large as their excitation potentials. Thus for sodium the ionization potential is 5·1 volts, for mercury 10·3 volts, *etc.*

In the case of the alkaline earths and some other elements, besides the ionization potentials in the narrower sense, experimental determinations have also been made of the *ionization potentials of the second order, i.e.*, those potentials at which the second most loosely bound electron is torn from the atom. For example, for magnesium the two potentials amount to 7·6 and 15 volts, and for helium to 24·5 and 78·6 volts.

Since the excitation potential is always *smaller* than the ionization potential, we must assume that, in the "*excited*" state, an atom still retains all its electrons, but that one electron is *more loosely bound than in the normal state.* The existence of more than a single *abnormal* state for the atoms of an element appears from the fact that, for the alkaline earths and for zinc, cadmium, and mercury, *two* characteristic potentials below the ionization potential can be sharply distinguished; *e.g.*, for magnesium at 2·7 and 4·3 volts, whereas the first ionization potential is 7·6 volts. In the case of mercury the corresponding three values are 4·9, 6·7, and 10·3 volts. In favourable conditions *Franck* and *Einsporn* [1] were able to detect three further characteristic potentials for mercury, at 5·5, 7·7, and 8·6 volts.

Observations on electron impacts lead to the conclusion that for every atom of an element, in addition to its normal state, a *series of higher states* is possible. Each of these higher states is characterized by a quite definite value of energy, which indicates the *excess of atomic energy* in that state over that in the normal state. Combining this result

[1] J. Franck and E. Einsporn, *Zeitschr. f. Phys.*, 2 (1920), p. 18.

with the hypothesis of *light-quanta, viz.*, with the assumption that the energy liberated by a *single radiating atom* is transformed into a *single light-quantum*, we come to the inevitable conclusion that the atoms of an element can *emit* light only of *quite definite frequencies characteristic* of the element. Furthermore, it follows that the manifold of these frequency values may be represented as *differences* between pairs of members of a *series* which is throughout *proportional* to the characteristic potentials.

In point of fact, as is well known, *Kirchhoff* and *Bunsen* in 1860 made the important discovery that, in the *spectra* of flames into which compounds of different *metals* have been introduced, *spectral lines* of definite wave-lengths are to be observed. These lines are characteristic of the metals, and their wave-lengths prove to be *independent* both of pressure and temperature, and, above all, of the particular compound of the metal used, so that they are to be regarded as an inherent property of *atoms* as such. In 1908 *Ritz* discovered, in connection with some older researches of *Rydberg's*, that the frequencies of spectral lines can be represented as *differences* between so-called *spectral terms*, which themselves form a much simpler manifold than the actual frequency values. The possibility of representing lines by terms led to the *combination principle* propounded by Ritz, according to which the frequency of a spectral line may be equated to the sum or difference of the frequencies of two other lines of the same element. The principle of representation in terms, which was first established only for *optical* spectra, found an important and accurate extension in its applicability to *X-ray spectra* on the subsequent discoveries of these.

The hypothesis of light-quanta leads us to suspect a connection between *excitation energy* and that *spectral line* which exhibits the *greatest intensity* in the spectrum of an element. In the case of *sodium*, for example, this is the well-known strong *yellow* line,[2] with a wave-length of $5{\cdot}89 \times 10^{-5}$ cm., and hence to be associated with an energy of $3{\cdot}34 \times 10^{-12}$ ergs (according to the formula $E = h\,c/\lambda$). This value, however, agrees very well with the previously quoted figure of $3{\cdot}37 \times 10^{-12}$ ergs, obtained from the *observed excitation potential* of sodium. Quite generally we must have

$$\frac{h\,c}{\lambda} = e\,V\,. \tag{2}$$

[2] The yellow sodium line really represents a doublet, as we shall see later.

Measuring the potential in volts instead of in the 300 times larger absolute units, and the wave-length (as is usual) in Ångström units ($10^{-8}$ cm.), we find

$$\lambda V = 1{\cdot}234 \times 10^4 \quad (\lambda \text{ in Å., } V \text{ in volts}). \tag{3}$$

This equation holds good generally for the relation between excitation potential and the strongest spectral line, and in the case of the alkaline earths, of zinc, cadmium, and mercury it is doubly satisfied, for the two strongest lines correspond to the two values of the characteristic potential. A similar equation is also found to apply to the *ionization potential*. For experience has shown that the spectral lines peculiar to an element can always be arranged in a number of so-called *series*, in each of which the lines *converge* in the direction of decreasing wave-length. In this direction the interval between two successive lines belonging to the same series becomes ever smaller, with a concurrent decrease in the sharpness of the lines themselves. To every series a *series limit* may thus be fixed, characterized by a definite frequency or wave-length. It is found now by experiment that the relation between the *ionization potential* and the limiting wave-length of a definite series (as we shall discuss more fully in a subsequent chapter) is actually represented by eqn. (3).

In spectroscopy it is customary to express *terms* in *reciprocal wave-lengths—wave-numbers*—so that a spectral line arising from the combination of two terms $T'$ and $T''$ is governed by the general relation

$$\frac{\nu}{c} = T' - T''. \tag{4}$$

The *numerically largest term* in the spectrum of an atom is called the *fundamental term*; *e.g.*, for sodium its value is 41,449, corresponding to a limiting wave-length of $2{\cdot}413 \times 10^{-5}$ cm. From what has been said above, the fundamental term of a neutral atom appears as identical with the ionization energy divided by $c\,h$.[3]

Thus, quite generally, the spectral terms multiplied by $c\,h$ may be regarded as the *energy-steps* or *energy-levels* corresponding to individual discrete states of the atom. The fundamental term corresponds to the *normal, fundamental, state*, whereas the numerically smaller terms are to

[3] To ionization potentials of higher orders correspond the fundamental terms in the so-called spark spectra, which are produced by previously ionized atoms (*vide* the subsequent chapter on atomic spectra).

be associated with states of greater energy, and so "*higher.*" For the "higher" the state, the smaller is the work (proportional to the term) required to remove an electron from that state. Conversely, the work necessary to raise an atom from its normal state to one characterized by the term $T'$ is given by

$$A = c\,h\,(T_0 - T'), \tag{5}$$

where $T_0$ denotes the fundamental term. Potentials may be transformed into term values with the help of the relation

$$e\,V = c\,h\,T, \tag{6}$$

so that, measuring in volts and hence multiplying the right-hand side by 300, we have

$$T = 8{\cdot}11 \times 10^3 \text{ cm.}^{-1} \text{ per volt.} \tag{7}$$

A transition between two different energy-levels is linked with a wave-length of which the reciprocal value is equal to the *difference* between the two corresponding terms. According as the initial or the final level is the higher, an *emission* or an *absorption* line is produced. The identity of the frequencies of the *absorption spectrum* with those of the emission spectrum, as *Kirchhoff* and *Bunsen* first discovered, is a necessary consequence of the concept of *energy-levels* in atoms based on the principle of light-quanta.

### § 94. The Sequence of the Elements.

It follows from the hypothesis of the nuclear atom that every *atomic species* must be *characterized by two integers*. One of these gives the number of *protons*, and, inasmuch as the mass of the electrons is negligible in comparison with that of the protons, it represents essentially the *atomic weight*. The second number states the *charge of the atomic nucleus*, referred to the elementary quantum of electricity as unit, and hence also the *number of planetary electrons* in the *neutral* state of the atom.

An approximate determination of this so-called *nuclear charge number* has been rendered possible by researches on the *passage of α-rays through metal foils*. As we have already mentioned in an earlier section (§ 89), $\alpha$-rays are deflected by the atomic nucleus in their passage through matter. The so-called *scattering* of the $\alpha$-rays thus produced therefore depends on the nuclear charge number of the metal traversed, and in this way this number has been calculated with an accuracy of 1%. By such a method *Chadwick*

in 1920 obtained for platinum a value of 77·4 (or, within the limits of error of 1%, 77 or 78), for silver 46·3, and for copper 29·3. The nuclear charge numbers are therefore considerably smaller than the known atomic weights of these three metals (193·1, 107·9, and 63·6 respectively), so that we may conclude that the atomic nuclei not only are surrounded by electrons, but also contain electrons in addition to protons.

A much more perfect method for the determination of nuclear charge numbers is afforded by the *principles of X-ray spectra*, as was discovered in 1913 by *Moseley*, the founder of X-ray spectroscopy. X-ray spectra, which we shall discuss in a subsequent chapter, exhibit series of lines that are of the same type for the various elements. As now Moseley discovered, the *elements* can be arranged in a *consecutively numbered series*, in such a way that the *square root of the frequency of a definite X-ray line increases linearly with the number of the element*.

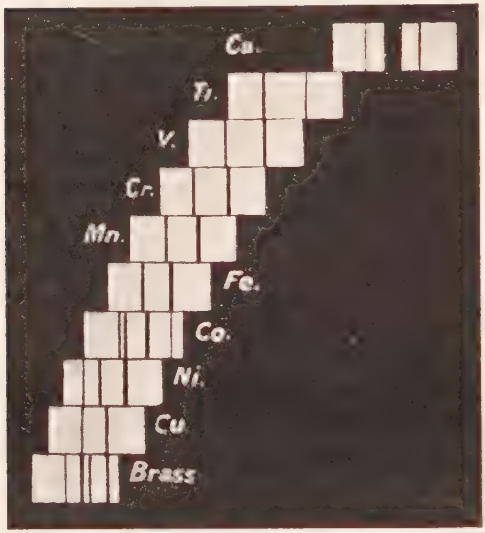



FIG. 61.—*K*-series (after Moseley).

The "*natural*" *sequence of the elements* thus obtained agrees on the whole, as Moseley recognized, with the series formed by arranging the elements in order of *increasing atomic weight*. Whereas, however, no regular relation exists between the atomic weight numbers, in the natural sequence the displacement of the X-ray lines from element to element takes place with the greatest exactness, so that the existence of any gap in the series is clearly revealed by too large a jump.

In Fig. 61, some photographs taken by Moseley of the so-called *K*-series for a sequence of elements are reproduced.

In addition to the strong $K_\alpha$- line, a second, weaker, ($K_\beta$-) line can be recognized. (It was not until spectroscopic methods had been perfected that further lines were first observed.) The photographs are so arranged that points lying vertically with respect to each other correspond to the same wave-length. The wave-lengths increase from left to right, and the frequencies from right to left, so that, in Fig. 61, the element with the lower number lies above the element with the higher number. The photographs refer to the elements calcium, titanium, vanadium, chromium, manganese, iron, cobalt, nickel, copper, and zinc : zinc was used alloyed with copper, in the form of brass.[1] They show with wonderful clearness that the given elements follow the natural sequence, but that there is an element missing between calcium and titanium. (This is scandium, which at the time was not available to Moseley, because of its rarity.) It may also be noticed—and this is most remarkable—that cobalt precedes nickel in the natural sequence, although it has the higher atomic weight of 58·97 as against 58·68. Brass shows both the zinc and the copper lines, and cobalt appears to be contaminated by nickel and iron.

Direct X-ray spectroscopic measurements, carried out with the help of crystals (as will be more fully described in a later section) have made it possible to establish the natural sequence of the elements from sodium to the highest element, *uranium*. Furthermore, indirect measurements of characteristic wave-lengths in the boundary region between X-rays and ultra-violet rays have enabled us to follow the course of the spectra from sodium down to the element with the second smallest atomic weight, helium.

The natural sequence, which undoubtedly commences with hydrogen in the first and helium in the second place, is thus *unequivocally* established. Among all known elements *uranium* has the highest nuclear charge number, *viz.*, 92. In our present state of knowledge, there are only *two gaps* in the sequence of the elements, corresponding to the numbers 85 and 87.[2] The series of the elements is represented in Table III, in which for each element the conventional name, the chemical symbol, and the atomic weight are given. This table, it is true, includes elements

[1] The alloy was used because zinc is not sufficiently resistant to heat.

[2] At the time of Moseley's discovery, there were still six gaps. In 1922 Hevesy and Coster discovered element No. 72 (hafnium), in 1925 Noddack and Take Nos. 43 and 75 (masurium and rhenium), and finally in 1926 Hopkins discovered No. 61 (illinium).

## Table III

*The Natural Sequence of the Elements*

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1 | Hydrogen | H | 1·0078 | 47 | Silver | Ag | 107·88 |
| 2 | Helium | He | 4·002 | 48 | Cadmium | Cd | 112·41 |
| 3 | Lithium | Li | 6·940 | 49 | Indium | In | 114·8 |
| 4 | Beryllium | Be | 9·02 | 50 | Tin | Sn | 118·7 |
| 5 | Boron | B | 10·82 | 51 | Antimony | Sb | 121·8 |
| 6 | Carbon | C | 12·000 | 52 | Tellurium | Te | 127·5 |
| 7 | Nitrogen | N | 14·008 | 53 | Iodine | I | 126·93 |
| 8 | Oxygen | O | 16·000 | 54 | Xenon | Xe | 130·2 |
| 9 | Fluorine | F | 19·00 | 55 | Cæsium | Cs | 132·8 |
| 10 | Neon | Ne | 20·18 | 56 | Barium | Ba | 137·36 |
| 11 | Sodium | Na | 23·00 | 57 | Lanthanum | La | 138·9 |
| 12 | Magnesium | Mg | 24·32 | 58 | Cerium | Ce | 140·13 |
| 13 | Aluminium | Al | 26·97 | 59 | Praseodymium | Pr | 140·9 |
| 14 | Silicon | Si | 28·06 | 60 | Neodymium | Nd | 144·3 |
| 15 | Phosphorus | P | 31·02 | 61 | Illinium | Il | — |
| 16 | Sulphur | S | 32·06 | 62 | Samarium | Sm | 150·43 |
| 17 | Chlorine | Cl | 35·46 | 63 | Europium | Eu | 152·0 |
| 18 | Argon | A | 39·94 | 64 | Gadolinium | Gd | 157·3 |
| 19 | Potassium | K | 39·10 | 65 | Terbium | Tb | 159·2 |
| 20 | Calcium | Ca | 40·07 | 66 | Dysprosium | Dy | 162·46 |
| 21 | Scandium | Sc | 45·10 | 67 | Holmium | Ho | 163·5 |
| 22 | Titanium | Ti | 47·90 | 68 | Erbium | Er | 167·64 |
| 23 | Vanadium | V | 50·95 | 69 | Thulium | Tm | 169·4 |
| 24 | Chromium | Cr | 52·01 | 70 | Ytterbium | Yb | 173·5 |
| 25 | Manganese | Mn | 54·93 | 71 | Lutecium | Lu | 175·0 |
| 26 | Iron | Fe | 55·84 | 72 | Hafnium | Hf | 178·6 |
| 27 | Cobalt | Co | 58·94 | 73 | Tantalum | Ta | 181·5 |
| 28 | Nickel | Ni | 58·69 | 74 | Tungsten | W | 184·0 |
| 29 | Copper | Cu | 63·57 | 75 | Rhenium | Re | — |
| 30 | Zinc | Zn | 65·38 | 76 | Osmium | Os | 190·9 |
| 31 | Gallium | Ga | 69·72 | 77 | Iridium | Ir | 193·1 |
| 32 | Germanium | Ge | 72·60 | 78 | Platinum | Pt | 195·23 |
| 33 | Arsenic | As | 74·96 | 79 | Gold | Au | 197·2 |
| 34 | Selenium | Se | 79·2 | 80 | Mercury | Hg | 200·6 |
| 35 | Bromine | Br | 79·92 | 81 | Thallium | Tl | 204·4 |
| 36 | Krypton | Kr | 82·9 | 82 | Lead | Pb | 207·21 |
| 37 | Rubidium | Rb | 85·5 | 83 | Bismuth | Bi | 209·0 |
| 38 | Strontium | Sr | 87·6 | 84 | Polonium | Po | (210) |
| 39 | Yttrium | Y | 88·9 | 85 | — | — | — |
| 40 | Zirconium | Zr | 91·22 | 86 | Radon | Rn | 222 |
| 41 | Columbium | Cb | 93·5 | 87 | — | — | — |
| 42 | Molybdenum | Mo | 96·0 | 88 | Radium | Ra | 226·0 |
| 43 | Masurium | Ms | — | 89 | Actinium | Ac | (226) |
| 44 | Ruthenium | Ru | 101·7 | 90 | Thorium | Th | 232·1 |
| 45 | Rhodium | Rh | 102·9 | 91 | Protoactinium | Pa | (230) |
| 46 | Palladium | Pd | 106·7 | 92 | Uranium | U | 238·14 |

for which it has not yet been possible to make X-ray spectroscopic photographs, such as the rare gases, mercury,

radium, polonium, actinium, protoactinium; the positions of these elements follow on the one hand from the chemical periodicity (to be discussed later on), and on the other hand from their atomic weights. In the natural sequence, the *atomic weight* increases throughout, with only *four exceptions*. The heavier element argon precedes the lighter element potassium (Nos. 18 and 19); cobalt comes before nickel (Nos. 27 and 28) as already mentioned; tellurium occurs before iodine (Nos. 52 and 53); and, finally, thorium precedes protoactinium (Nos. 90 and 91). The nuclear charge numbers found on the basis of Moseley's relation agree, in fact, with those which were subsequently derived by Chadwick's method. The identity of the ordinal (or atomic) numbers of Moseley's sequence with the nuclear charge numbers was established by *van den Broek* in 1913.

## § 95. Chemical Periodicity.

In the natural sequence of the elements, experience shows that *six elements* occupy a *very special position*, while simple numerical relations exist between their atomic numbers. These elements are, in fact, absolutely *passive* from the *chemical* point of view; in contrast to all the other elements, they form *no chemical compounds* whatsoever. They are the so-called *rare gases*,[1] *viz.*, helium, neon, argon, krypton, xenon, and radium emanation (or radon).[2] Their atomic numbers are

$$2,\ 10,\ 18,\ 36,\ 54,\ 86.$$

The differences between these numbers are

$$8,\ 8,\ 18,\ 18,\ 32,$$

and we may suspect in advance a fundamental reason for the fact that the first atomic number and the following differences are of the form $2\,z^2$ ($z = 1, 2, 3, 4$).

From the point of view of modern atomic theory, *chemical activity* must obviously be referred to *electrical forces* pro-

[1] The rare gases (apart from radon) were discovered by *Ramsay* in the atmosphere towards the end of the nineteenth century.

[2] The "International" Committee on Chemical Elements has proposed the names *radon*, *thoron*, and *actinon* for the three radioactive emanations. *Radon* thus replaces *niton*, but the term *radium emanation* is also still widely used.

ceeding from the atoms. It is these forces which lead to a combination of oppositely charged atoms, and thus to the *formation of molecules*. Since, in the neutral atom, the number of electrons revolving around the nucleus is equal to the nuclear charge number, an atom can only become chemically active when the number of electrons surrounding its nucleus becomes either more or less than the nuclear charge number. In the first case, the atom as a whole becomes electrically negative ; in the second case, it becomes positive ; moreover, the charge is determined by the number of electrons gained or lost. It must therefore be concluded from the chemical passivity of the rare gases that, in their atoms, the *arrangement of the electrons about the nucleus* is generally *stable*, in the sense that electrons are neither removed from the atomic system, nor are new ones introduced into it—a state of affairs that appears possible in all other atoms.

Thus the rare gases to a certain extent represent landmarks in the natural sequence of the elements, and they divide the series into *seven periods*. Of these, the first period contains the first *two* elements (hydrogen and helium) ; the second and third periods each contain *eight* elements, Nos. 3 to 10 (lithium to neon), and Nos. 11 to 18 (sodium to argon). The fourth and fifth periods each comprise *eighteen* elements, Nos. 19 to 36 (potassium to krypton), and Nos. 37 to 54 (rubidium to xenon). The sixth period includes *thirty-two* elements, Nos. 55 to 86 (cæsium to radon). The seventh period, beginning with the unknown element No. 87, apparently breaks off in the sixth place at uranium ; and it must remain an open question whether elements with still higher nuclear charge are not capable of existence, have already disintegrated, or have not yet been discovered because of their rarity.[3]

As long ago as 1869, *Lothar Meyer* and *Mendeléeff* simultaneously, but independently, made the important discovery that, if the elements be arranged in order of increasing atomic weight, characteristic physical and chemical properties recur *periodically*, and that, if the successive periods be written down in successive horizontal rows, an arrangement is possible in which elements occurring in the same vertical column are distinguished by similar properties. The further development of the so-called *periodic system* resulted in a distinction between *eight vertical groups*, each with two subgroups. An exact representation, however, was not possible

[3] *Cf.* § 126.

## TABLE IV

*The Periodic System of the Elements*

| | I A | I B | II A | II B | III A | III B | IV A | IV B | V A | V B | VI A | VI B | VII A | VII B | VIII A | VIII B |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 H | | | | | | | | | | | | | | | 2 He |
| 2 | 3 Li | | 4 Be | | 5 B | | 6 C | | 7 N | | 8 O | | 9 F | | | 10 Ne |
| 3 | 11 Na | | 12 Mg | | 13 Al | | 14 Si | | 15 P | | 16 S | | 17 Cl | | | 18 A |
| 4 | 19 K | 29 Cu | 20 Ca | 30 Zn | 21 Sc | 31 Ga | 22 Ti | 32 Ge | 23 V | 33 As | 24 Cr | 34 Se | 25 Mn | 35 Br | 26 Fe 27 Co 28 Ni | 36 Kr |
| 5 | 37 Rb | 47 Ag | 38 Sr | 48 Cd | 39 Y | 49 In | 40 Zr | 50 Sn | 41 Cb | 51 Sb | 42 Mo | 52 Te | 43 Ms | 53 I | 44 Ru 45 Rh 46 Pd | 54 Xe |
| 6 | 55 Cs | 79 Au | 56 Ba | 80 Hg | 57—71 rare earths | 81 Tl | 72 Hf | 82 Pb | 73 Ta | 83 Bi | 74 W | 84 Po | 75 Re | 85 | 76 Os 77 Ir 78 Pt | 86 Rn |
| 7 | 87 | | 88 Ra | | 89 Ac | | 90 Th | | 91 Pa | | 92 U | | | | | |

until the knowledge of atomic numbers also led to an exact recognition of the number of places in the periods.

Experience shows that the table of the periodic system is to be constructed in the following manner (Table IV.). Eight vertical columns are formed, each with two sub-groups. The groups corresponding to the vertical columns are denoted by I, II, *etc.*, up to VIII; the sub-groups may be distinguished in the table by the letters A and B. The two elements in the first period are now placed in sub-groups I A and VIII B. The eight elements each of the second and the third period are to be allotted to the eight groups in such a way that the first two in each period fall in A columns, and the last six in B columns. In the fourth and fifth periods, each containing 18 elements, we proceed as follows: we place the first eight elements in the eight A sub-groups, and the last eight in the eight B sub-groups, whilst the ninth and tenth elements are allotted to sub-group VIII A, so that, in each of the fourth and fifth periods, this sub-group contains a *triad*. In the sixth period we follow the same arrangement as in the fourth and fifth, but with the difference that in sub-group III A a *group of fifteen elements* is to be placed, instead of a single element. Thus, in the sixth period, one of the 16 positions in the eight A and eight B sub-groups will contain 3 elements, and another 15 elements. Finally, in the seventh period, the six elements are to be placed in the A sub-group. The three triads consist of the *iron*, the *palladium*, and the *platinum metals*. The group of fifteen elements is known as the *rare earth* group.

## Chapter XII

# THE PRINCIPLES OF ATOMIC MECHANICS

### § 96. De Broglie's Theory.

The hypothesis of *material waves* leads to a result of particular importance, as *de Broglie* recognized, when it is applied to the *periodic motions of a material particle.* We know the propagation of a wave to be characterized by the existence, at the same instant, of the *same condition of vibration* at all points *along a ray* which are at a *distance* of an *integral multiple of the wave-length* from one another. Since it is impossible for two different conditions of vibration to be simultaneously present at one and the same place within a wave, a *movement in a closed path at constant velocity*, in conformity with the concept of material waves, appears to be possible only if the *length of path* ($l$) be an *integral multiple of the wave-length, i.e.*, if

$$\frac{l}{\lambda} = n \quad (n = 1, 2, 3 \text{ } etc.). \tag{1}$$

If the *velocity*, and consequently the wave-length, be *variable*, we must obviously resolve the length of path $l$ into elements $ds$ and divide each by the corresponding value of the wave-length. Denoting the integral taken over the closed path by the usual symbol $\oint$, we obtain in place of eqn. (1) the more general relation

$$\oint \frac{ds}{\lambda} = n. \tag{2}$$

In accordance with § 91 eqn. (2) we may write this

$$\oint G\, ds = n\, h; \tag{3}$$

*i.e.*, the line-integral of the momentum taken over the closed orbit represents an integral multiple of the elementary quantum of action. In the special case of a *circular motion* proceeding with *constant velocity*, we have thus the product

of mass, velocity, and radius of the orbit—the *angular momentum*—

$$m\,a\,v = \frac{n\,h}{2\,\pi}. \tag{4}$$

The concept of material waves thus leads necessarily to a conclusion which was propounded in 1913 by *Bohr* as an independent hypothesis, and was applied by him with very great success to the interpretation of line-spectra. Bohr equated the angular momentum for the *normal state of the hydrogen atom* multiplied by $2\pi$ to the elementary quantum of action, while for the abnormal states he assumed this product to be equal to $2h$, $3h$, $4h$, *etc.* Thus, first he ascertained from the principles of classical mechanics the totality of all possible orbits in the atom, and then he *singled out* from this totality as "*admissible by the quantum theory*" those orbits which satisfy the *quantum condition* expressed in eqn. (4). This selection of "quantized" orbits inevitably remained entirely unintelligible from the point of view of classical mechanics, as also did the determination of mechanical orbits by the optical fundamental constant $h$. Both difficulties were first solved by de Broglie's theory.

The *velocity of propagation of a material wave* is, according to a generally valid relation of wave theory, equal to the product of wave-length and frequency; *i.e.*,

$$u = \lambda\,\nu. \tag{5}$$

Introducing for $\lambda$ the value given in § 91 eqn. (12)—$h/m\,v$—and similarly replacing $h\,\nu$ by $m\,c^2$, we obtain the relation

$$u = \frac{c^2}{v}. \tag{6}$$

The velocity of material waves is therefore *inversely proportional* to the corresponding mechanical velocity.

This peculiar relation of de Broglie's theory establishes a close bond between two old physical principles, one of which was regarded as a supreme *geometrical-optical* theorem, and the other as a supreme *mechanical* theorem. The optical theorem was propounded in the first half of the seventeenth century by *Fermat* as the principle of *least time*. According to this principle, the *actual path* traversed by light between two points is such that the time taken is a *minimum*.[1] Since the time is equal to the quotient of

[1] Fermat's principle enables the path of a beam of light to be accurately determined; *e.g.*, the law of refraction can very easily be derived from it.

the path and the velocity, Fermat's principle is mathematically expressed by the postulate that the integral along the path of the reciprocal wave velocity $\left(\int ds/u\right)$ shall be a minimum.

The supreme *mechanical* principle, which was established by *Maupertuis* in the middle of the eighteenth century, and was subsequently formulated more accurately by *Euler*, postulates that the integral of the mechanical velocity taken along the path $\left(\int v\, ds\right)$ shall be a minimum.[2] In Maupertuis's theorem *mechanical velocity* therefore plays the same part as *reciprocal wave velocity* in Fermat's theorem. Now, on certain conditions,[3] the latter is an approximately valid general principle for *all* kinds of waves. Hence it follows from the concept of material waves [*cf.* eqn. (6)] that the *mechanical principle* of Maupertuis is a *necessary consequence* of the general *wave principle* of Fermat.

Since the mechanical velocity $v$ is in any case less than $c$. the *material waves* associated, according to eqn. (6), with a moving body must travel with *velocities larger than that of light*. For a fast motor car the waves would travel some $10^{14}$ times more rapidly than the car itself, and for slower motions the disproportion would be still greater. This apparent contradiction, however, finds a satisfactory solution in the concept of *group velocities*, which was introduced into physics in the middle of the nineteenth century. By a *group of waves* we mean a set of waves the frequencies of which lie within certain limits. The behaviour of such a group becomes intelligible when we remember that the mean value of the sine of all possible angles vanishes, if the sign be taken into account. To every positive value of the sine corresponds an equal but negative value, and conversely, so that there is a mutual cancellation of values. If, therefore, we consider a set of vibrations of which the phases do not in any way agree, the resultant vibration, which is determined by the sum of the products of the amplitudes

[2] *Cf.* the author's "*Grundgleichungen der Mechanik*" (Leipzig; Veit, 1914), lect. 19. More general and, from the point of view of the theory of relativity, more accurate is the condition that $\int G\, ds$ shall be a minimum.

[3] The principal of these conditions is that the linear dimensions of the region to which the theorem is applied shall be large in comparison with the wave-length.

and the sines of the corresponding phases, will in general be very small in comparison with the algebraical sum of the amplitudes.

Let us now consider a point within a group of waves at which there is a fairly extensive agreement in phase between the individual waves belonging to the group; the resultant vibration will then be proportionately large. It is obviously characteristic of such a point (which may be called an *energy centre* of the group of waves) that, for two waves of which the frequencies differ infinitesimally, the difference in phase will be vanishingly small and of the second order. Denoting the phase by $\psi$, we have therefore for an energy centre

$$\frac{\partial \psi}{\partial \nu} = 0. \tag{7}$$

Now, as we know from § 42 eqn. (7), the *phase* is determined by the relation

$$\psi = 2\pi\nu\left(t - \frac{s}{u}\right) + \epsilon, \tag{8}$$

where $t$ is the time, $\epsilon$ the phase-constant, and $s$ the distance from the source. Hence

$$\frac{\partial \psi}{\partial \nu} = 2\pi t - 2\pi s \frac{d\left(\frac{\nu}{u}\right)}{d\nu} + \frac{d\epsilon}{d\nu}, \tag{9}$$

and from eqn. (7) we obtain

$$s = \left[\frac{d\left(\frac{\nu}{u}\right)}{d\nu}\right]^{-1}\left(t + \frac{1}{2\pi}\frac{d\epsilon}{d\nu}\right). \tag{10}$$

Since the quantity $s$ thus depends on the time, the energy centre itself must move, its velocity of motion being given by

$$g = \left[\frac{d\left(\frac{\nu}{u}\right)}{d\nu}\right]^{-1}, \tag{11}$$

and therefore the quantity $g$ represents the velocity of propagation of the energy of the waves. It is known as the *group velocity*, a term due to *Lord Rayleigh*, and is to be

distinguished from the wave velocity except in the special case that the latter is independent of the frequency.[4]

Now, according to eqn. (6), we have for material waves

$$\frac{\nu}{u} = \frac{\nu\, v}{c^2}, \tag{12}$$

and therefore

$$\frac{d\left(\frac{\nu}{u}\right)}{d\nu} = \frac{v}{c^2} + \frac{\nu}{c^2}\frac{dv}{d\nu}. \tag{13}$$

Again

$$\nu = \frac{m\, c^2}{h}, \tag{14}$$

and taking into account § 88 eqn. (6), which governs the relation between the mass and the stationary mass ($m_0$) and $v$, we find

$$\nu = \frac{m_0\, c^2}{h\sqrt{1 - \frac{v^2}{c^2}}}. \tag{15}$$

From this it follows that

$$\frac{dv}{d\nu} = \frac{m_0\, v}{h\left(1 - \frac{v^2}{c^2}\right)^{\frac{3}{2}}} = \frac{\nu\, v}{c^2 - v^2}. \tag{16}$$

Introducing this value into eqn. (13), we find

$$\frac{d\left(\frac{\nu}{u}\right)}{d\nu} = \frac{v}{c^2} + \frac{c^2 - v^2}{c^2\, v} = \frac{1}{v}, \tag{17}$$

and a comparison with eqn. (11) yields the important result

$$g = v. \tag{18}$$

The *mechanical velocities of material particles* therefore prove to be *identical with the group velocities of material waves.*

Conversely, as the author of this book has shown,[5] we

[4] Effecting the differentiation in eqn. (11), we find

$$\frac{1}{g} = \frac{1}{u} - \frac{\nu}{u^2}\frac{du}{d\nu}.$$

Hence, in the absence of dispersion, *i.e.*, when $u$ is independent of $\nu$, group and wave velocities are identical.

[5] A. Haas, *Physikal Zeitschr.*, 28 (1927), p. 632. For this demonstration and for several details of de Broglie's theory, *cf.* the author's book "*Wave Mechanics and the New Quantum Theory*" (London: Constable, 1928).

can start from the hypothesis that mechanical velocities represent group velocities of waves of any kind, the energy of the waves being composed of elements $h\nu$, and we can then derive the *fundamental laws of the theory of relativity from classical mechanics.*[6]

## § 97. Schrödinger's Theory.

The concept of material waves suggests that we should choose, as the basis of atomic mechanics, a differential equation of a type to describe quite generally the propagation of waves of any kind whatsoever. As we know from elementary mathematical physics, this process may be most simply represented by a *wave equation* of the form [1]

$$\nabla^2 S + \frac{4\pi^2}{\lambda^2} S = 0, \tag{1}$$

where $S$ denotes the oscillating quantity, and $\lambda$ the wave-length.

Let us apply this equation to the *material waves* of a *single particle*, leaving the significance of the "field-scalar" $S$ undetermined for the time being. We have by § 91 eqn. (12)

$$\frac{1}{\lambda^2} = \frac{m^2 v^2}{h^2}. \tag{2}$$

Again, denoting the *total energy* of the particle by $E$ and its *potential energy* by $V$, we have [2]

$$m v^2 = 2 (E - V). \tag{3}$$

Eqn. (1) may therefore be written in the form

$$\nabla^2 S + \frac{8\pi^2 m}{h^2} (E - V) S = 0. \tag{4}$$

[6] These are the formulæ for the dependence of mass on velocity, and for longitudinal and transversal mass: the principle of the inertial mass of energy; and the so-called Lorentz-transformation of time.

[1] According to § 43 eqn. (1), the general differential equation of a wave is given by

$$\frac{\partial^2 S}{\partial t^2} = \frac{\lambda^2}{\tau^2} \nabla^2 S,$$

where $\tau$ represents the period. If, in particular, the periodic time function be a *sine function*, so that

$$S = A \sin\left(\frac{2\pi t}{\tau} + B\right),$$

where $B$ is a function of the coordinates, we find

$$\frac{\partial^2 S}{\partial t^2} = -\frac{4\pi^2}{\tau^2} S,$$

and eqn. (1) follows.

[2] For $(E - V)$ is equal to the kinetic energy $(\frac{1}{2} m v^2)$.

This differential equation was propounded in 1925 by *Schrödinger* as the *fundamental equation of atomic mechanics.* He sought thereby to extend to mechanics a generalization corresponding to the development of the coarser *geometrical optics* based on *rays* to the finer *physical optics* based on *waves.*

Now the *theory of differential equations* has long taught us that the possibility of finding a single-valued, finite, and continuous solution for a differential equation exists only for definite values of the *parameters* which occur in the equation: these values are known as the *proper values* for the differential equation. In Schrödinger's equation the total energy $E$ represents a constant parameter of this kind, whereas the potential energy depends on the coordinates. The proper values appertaining to Schrödinger's equation in any special case thus determine quite definite *discrete values of the energy.* In this way the physical problem of *quantizing energy* was *reduced* by Schrödinger to a *purely mathematical* one, *viz.*, to the problem of proper values in the theory of differential equations, and this has long been solved for all the more important cases.

If we consider, for example, the differential equation

$$\frac{d^2y}{dx^2} + (a - x^2)\, y = 0, \tag{5}$$

it follows from purely mathematical considerations that *all odd integers*, but *only* these (thus $a = 1, 3, 5,$ *etc.*), are proper values for this equation.[3] As a second example, let us consider the differential equation of *spherical functions*, which is obtained by *transforming* the Laplace derivative from orthogonal to *spherical polar coordinates.* These consist of the distance ($r$) from a fixed point and of two angles which, on the analogy of the earth, correspond to geographical longitude and the complement of the latitude. The transformation leads to an expression for the derivative which, in the special case that $r$ is assumed to be constant and equal to the unit of length, may be denoted by $\nabla'^2$. It follows then from the theory of spherical functions that the differential equation

$$\nabla'^2\psi + a\,\psi = 0 \tag{6}$$

possesses a finite and continuous solution on the whole surface of the sphere only when

[3] *Cf.*, for example, Courant and Hilbert, "*Methoden der mathematischen Physik,*" (Berlin; Springer, 1924), Vol. I., p. 261.

$$a = n\,(n+1) \tag{7}$$

for *integral values* of $n$.[4] The proper values for eqn. (6) are hence the numbers 2, 6, 12, 20, 30, 42, 56, *etc.*

The two examples drawn from the mathematical theory of differential equations permit of an immediate application to the two physically important problems of the linear oscillator and of the rotator. By an *oscillator* we mean a particle vibrating about a fixed centre, such as an electron. In the special case of the vibrations occurring along a straight line, the oscillator is termed *linear*, and, when we are dealing with pure sine vibrations, it is said to be *harmonic*. If $q$ be the distance from the centre, $\nu$ the frequency, and $A$ the amplitude, we have

$$q = A \sin 2\,\pi\,\nu\,t\,. \tag{8}$$

Differentiation with respect to the time gives the velocity, and hence we find for the kinetic energy as half the product of the mass and the square of the velocity

$$L = 2\,A^2\,\pi^2\,\nu^2\,m\,\cos^2 2\,\pi\,\nu\,t\,. \tag{9}$$

Since the largest possible value of $L$ must represent the total energy (for the potential energy vanishes when $q = 0$), we have

$$E = 2\,A^2\,\pi^2\,\nu^2\,m\,, \tag{10}$$

and we obtain for $V$, as the difference between $E$ and $L$,

$$V = 2\,\pi^2\,\nu^2\,m\,q^2\,. \tag{11}$$

In the special case of a linear harmonic oscillator, Schrödinger's equation therefore assumes the form

$$\frac{d^2S}{dq^2} + \frac{8\,\pi^2\,m}{h^2}\,(E - 2\,\pi^2\,\nu^2\,m\,q^2)\,S = 0\,. \tag{12}$$

We put now for brevity

$$\frac{8\,\pi^2\,m}{h^2}\,E = a \tag{13}$$

and

$$\frac{4\,\pi^2\,m\,\nu}{h} = b\,. \tag{14}$$

Introducing a new variable $x$ defined by the equation

$$x = q\,\sqrt{b}\,, \tag{15}$$

we have

[4] *Cf.* Courant-Hilbert, Vol. I., p. 264 and p. 422.

$$\frac{d^2S}{dq^2} = b\frac{d^2S}{dx^2} \tag{16}$$

and

$$q^2 = \frac{x^2}{b}. \tag{17}$$

Hence, on dividing eqn. (12) throughout by $b$, we find

$$\frac{d^2S}{dx^2} + \left(\frac{a}{b} - x^2\right) S = 0. \tag{18}$$

This equation is of exactly the same form as the previous eqn. (5), $a\,b$ and $S$ now replacing $a$ and $y$ respectively. Since odd integers only are proper values for this equation, we must have

$$\frac{a}{b} = 2\,n + 1. \tag{19}$$

If follows, however, from eqn. (13) and (14) that

$$\frac{a}{b} = \frac{2\,E}{h\,\nu}, \tag{20}$$

so that

$$E = (2\,n + 1)\,\frac{h\,\nu}{2}. \tag{21}$$

In Schrödinger's mechanics, therefore, the energy of a linear harmonic oscillator appears as an *odd multiple of half an energy element*, if by this we understand, as usual, the product $h\,\nu$. This result differs from that of the older quantum theory, according to which the energy had always to be equated to an *integral multiple* of a whole energy element. In contradistinction to the older theory, it follows from Schrödinger's mechanics that the *lowest value* which the *energy* of an oscillator can assume is *other than zero*. Now theoretical physics refers the *heat content of solid bodies* essentially to oscillations executed by the atoms of the body about certain fixed centres, the energy of these oscillations manifesting itself as heat. Several phenomena observed at extremely low temperatures had led some physicists—especially *Nernst*—even before Schrödinger to the hypothesis of the so-called *energy of the absolute zero*. According to this hypothesis, even at the absolute zero, at which the older theory demanded the complete disappearance of vibrational energy, every oscillator would still possess energy to the amount of $h\,\nu/2$. This assumption, which had previously seemed very artificial, follows

naturally from the fundamental relation of Schrödinger's theory.

Like the oscillator, the *rotator* also played an important part in the older quantum theory, and here again Schrödinger's theory leads to results different from those originally obtained. By a rotator we mean a particle which moves at a *constant distance about a fixed centre of attraction*. If the motion be restricted to a definite plane, so that the particle moves in a circular orbit, we speak of a rotator with a *fixed axis*. If this restriction be removed, so that the particle travels on a *spherical surface*, we speak of a rotator with a *free axis*. We shall discuss the latter kind in the following considerations.

When the fixed distance is equal to the unit of length, the Laplace derivative in Schrödinger's differential equation may be replaced by the previously used symbol $\nabla'^2$. Inasmuch as $r$ is equal to unity, and the Laplace derivative contains the squares of the differentials of the coordinates in the denominators, we may divide by $r^2$, and thus replace $\nabla^2 S$ by $\nabla'^2 S/r^2$. Since the potential energy of a rotator is constant, so also is its kinetic energy ($L$), and, after multiplying throughout by $r^2$, we may therefore write Schrödinger's equation in the form

$$\nabla'^2 S + \frac{8\pi^2 m}{h^2} r^2 L S = 0. \tag{22}$$

In accordance with eqns. (6) and (7), the proper values for this equation are determined by the condition that, for *integral values* of $n$,

$$\frac{8\pi^2 m r^2 L}{h^2} = n(n+1). \tag{23}$$

Introducing, now, the *angular momentum* $U$ $(= m v r)$ we have

$$L = \tfrac{1}{2} m v^2 = \frac{U^2}{2 m r^2}. \tag{24}$$

Hence

$$\frac{4\pi^2 U^2}{h^2} = n(n+1),$$

or,

$$U = \frac{h}{2\pi}\sqrt{n(n+1)}. \tag{25}$$

Thus, for a rotator also, Schrödinger's theory leads to a *different* conclusion from that derived from the older quan-

tum theory, which equated the angular momentum to an integral multiple of $h/2\pi$.

A *generalization of Schrödinger's differential equation* is necessary when we consider processes in which more than a single material particle is concerned. If we denote by $\nabla_1^2$ the sum of the second partial derivatives of the first particle with respect to the coordinates, and by $m_1$ its mass, *etc.*, Schrödinger's equation for the *problem of several particles* assumes the form

$$(26)\quad \frac{1}{m_1}\nabla_1^2 S + \frac{1}{m_2}\nabla_2^2 S + \ldots + \frac{8\pi^2}{h^2}(E - V)S = 0,$$

as will be established in the following section. It must be observed, however, that this equation is no longer descriptive of a true wave process in the ordinary sense. The process described by it can be regarded as a "*wave*" only in a so-called *configurational space*, the *dimensions* of which are *thrice the number of particles*, whilst the coordinates are to be interpreted as intervals multiplied by the square root of the mass of the particle in question.[5]

### § 98. The Quantum Mechanics of Heisenberg.

Schrödinger's theory, based on the concept of material waves, does not represent the only attempt to rebuild atomic mechanics. It was preceded by a few months by the *quantum mechanics* created by *Heisenberg*, although, as Schrödinger himself demonstrated subsequently, the two theories are intimately connected. This is all the more remarkable because the object to which Heisenberg devoted himself in developing his theory differed completely from the goals at which de Broglie and Schrödinger aimed. Both these investigators were endeavouring to explain atomic processes with the help of the methods of classical physics, and so to bridge the gulf separating classical from "quantized" atomic physics. Heisenberg, however, started with quite the opposite conception, *viz.*, that only a complete rejection of the customary classical ideas could lead to a proper understanding of atomic principles. He was convinced that any attempt to form general views in atomic theory must be avoided, as a matter of course, as aimless and of no meaning, and that all quantities *not accessible to*

[5] The complication of ascribing to the coordinates the dimensions of length multiplied by the square root of the mass is indeed, avoidable when all the particles concerned have the same mass.

*experimental observation* must be *expelled* from theoretical physics.

In place of the earlier quantum theory, which made use of the concept of *electronic orbits*—a concept of which the accuracy could not be checked—Heisenberg thus sought to establish a new quantum mechanics which should deal solely with relations between observable quantities in atomic physics. As such quantities he regarded only the *frequencies* and *intensities* of the *spectral lines* characteristic of the atoms, in addition to the *energy-levels in the atoms*, which could be determined from experiments on electron collisions. In his view the problem of the new atomic mechanics was two-fold. First, numerical quantities had to be constructed out of these spectral quantities to take the place of the coordinates and velocities of the electrons. Secondly, these numerical quantities were to be combined, as far as possible along the lines of classical mechanics, into relations which should lead to the establishment of rules for the application of quantum principles.

For this purpose Heisenberg replaced the *periodically variable coordinates of the electrons* by a *set of partial vibrations*, the *frequencies* of which should agree with those of the *spectral lines* generated by the electron concerned in any of its transitions. Since each spectral frequency is associated with a *transition between two atomic states*, the partial vibrations executed by an individual electron must form a *double manifold*,[1] and thus may be represented by a so-called quadratic scheme. If we allot zero to the normal state, and number the other states consecutively, this scheme may be given the form

$$\begin{array}{cccc} q_{00} & q_{01} & q_{02} & q_{03} \cdots \\ q_{10} & q_{11} & q_{12} & q_{13} \cdots \\ q_{20} & q_{21} & q_{22} & q_{23} \cdots \\ q_{30} & q_{31} & q_{32} & q_{33} \cdots \\ \vdots & \vdots & \vdots & \vdots \end{array}$$

In subsequent considerations the subscripts may be suitably written in brackets: $q(n, m)$ then denotes a vibration occurring with a frequency $\nu(n, m)$ and with an amplitude which in itself represents a measure of the *probability of a transition* between the $m$th and the $n$th states. The square of this

[1] At the beginning of the nineteenth century Fourier showed, in his well-known theorem, that any quite arbitrary periodic variability can always be represented by a set of partial vibrations of pure sine form. Heisenberg adopted the principle of Fourier's theorem, replacing the simple Fourier series by a double manifold.

amplitude shall determine the *intensity of the spectral line* corresponding to the transition between those two states.

A scheme of the kind represented above is known in algebra as a *matrix*. The calculus of matrices was developed in the middle of the nineteenth century, various operations in calculation being suitably defined. Thus the *product* of two matrices $A \times B$ is a matrix $C$, where

$$C(n, m) = \sum_k A(n, k)\, B(k, m). \tag{1}$$

It is easy to see from this definition that the product $B\,A$ must in general differ from the product $A\,B$, *i.e.*, that the *multiplication of matrices cannot be* generally *commutative.*

Heisenberg regarded coordinates and momenta as numerical quantities such as we have described, and he based his new quantum mechanics upon the hypothesis that, for the multiplication of a momentum component of an electron by the corresponding coordinate, the so-called *commutation rule* holds good :—

$$p\,q - q\,p = \frac{h}{2\pi i}\,\epsilon\,. \tag{2}$$

In this equation $h$ denotes the elementary quantum of action, $i$ (as usual) $\sqrt{-1}$, and $\epsilon$ a so called *unit-matrix* of the form

$$\begin{matrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{matrix} \textit{ etc.}$$

According to eqn. (2) the error, occasioned by regarding momentum and coordinate as ordinarily interchangeable quantities in the multiplication, would be of the order of absolute magnitude $10^{-27}$, and therefore quite negligible in macromechanics.

*Born* and *Jordan*, to whom the development of Heisenberg's ideas into a systematic *matrix mechanics* is due,[2] have been able to show that the application of the commutation rule to matrices formed from spectral quantities yields both the *Law of Conservation of Energy* and the *principle of light-quanta* as necessary consequences.[3] In its application to oscillators and rotators, the results of matrix mechanics agree with those of the mechanics of proper

[2] Although Heisenberg himself established the rule of multiplication expressed in eqn. (1), he did not recognize its connection with the calculus of matrices.

[3] *Cf.* the author's "*Wave Mechanics and the New Quantum Theory*," Ch. ix.

values, as they also do in numerous other applications which we shall discuss in subsequent sections.

The *mathematical equivalence* of the Heisenberg-Born-Jordan quantum mechanics with Schrödinger's mechanics of proper values, despite the complete difference between the underlying hypotheses, was first demonstrated by Schrödinger himself.[4] In order to understand the fundamental ideas of his proof, we shall start from the concept of the *mathematical operator*. As we know, we can symbolically represent the differentiation of a function $y$ with respect to an argument $x$ as the multiplication of $y$ by an operator $d/dx$. Similarly, the second partial derivative of $y$ with respect to $x$ may be regarded as the multiplication of $y$ by an operator $\partial^2/\partial x^2$. Again, we may take $x$ itself as an operator, thereby meaning a multiplication by $x$; and in a like sense we may speak of an operator 1 which, when applied to any function, gives the function itself. By the *product of two operators* we shall understand an operator which has the same effect as the successive application of those two operators. Thus the second derivative of a function appears as the *square* of the first, and the operator product $\frac{d}{dx}\,x$ signifies the differentiation with respect to $x$ of the product of a function and $x$.

It follows from the above definition that the *multiplication* of two operators is *not commutative*, so that *commutation rules* can be established also for operators. One of these may be derived from the identity

$$\frac{\partial(x\,y)}{\partial x} = y + x\frac{\partial y}{\partial x}. \tag{3}$$

Limiting ourselves to the operators by omitting the function $y$, we obtain the operator equation

$$\frac{\partial}{\partial x}x - x\frac{\partial}{\partial x} = 1; \tag{4}$$

*i.e.*, if we multiply any function by $x$ and differentiate the product according to $x$, and if we then subtract from the result the product of $x$ and the derivative of the function with respect to $x$, we obtain the original function.

Let us now consider, instead of the operator $x$, the

[4] E. Schrödinger, "*Über das Verhältnis der Heisenberg-Born-Jordanschen Quantenmechanik zu der meinen*" [*Ann. der Phys.*, **79** (1926), p. 734]. *Vide* also his "*Collected Papers on Wave Mechanics*" (London: Blackie, 1928).

operator $q$ in particular, meaning thereby a multiplication by the coordinate $q$. If, furthermore, we multiply eqn. (4) throughout by the expression $h/2\pi i$, we find a *commutation rule* of the form

$$\frac{h}{2\pi i}\frac{\partial}{\partial q}q - q\frac{h}{2\pi i}\frac{\partial}{\partial q} = \frac{h}{2\pi i}1. \tag{5}$$

Comparing eqns. (2) and (5), we see that, if the coordinate matrix be replaced by the operator $q$, the *momentum matrix* must be *replaced by the operator*

$$\frac{h}{2\pi i}\frac{\partial}{\partial q}. \tag{6}$$

We shall now attempt a generalization of mechanics by substituting for each coordinate, in the equation expressing the *conservation of energy* for a single particle, the corresponding operator $q$, and for each component of the momentum the corresponding expression (6). The kinetic energy is given by [5]

$$L = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2). \tag{7}$$

The Law of the Conservation of Energy may therefore be expressed by the equation

$$\frac{1}{2m}(p_x^2 + p_y^2 + p_z^2) + V - E = 0, \tag{8}$$

where $E$ denotes the constant total energy, and $V$ the potential energy. Regarding the left-hand side of this equation as an operator, we have to replace the expression in brackets by

$$\left(\frac{h}{2\pi i}\right)^2\left(\frac{\partial^2}{\partial q_x^2} + \frac{\partial^2}{\partial q_y^2} + \frac{\partial^2}{\partial q_z^2}\right), \text{ or } -\frac{h^2}{4\pi^2}\nabla^2, \tag{9}$$

in accordance with formula (6).

Thus, if eqn. (8) be regarded as an operator equation and we apply it to an arbitrary function $S$, we find the relation

$$\left\{-\frac{h^2}{8\pi^2 m}\nabla^2 + V - E\right\}S = 0, \tag{10}$$

or,

[5] For, denoting the velocity components by $v_x$, $v_y$, $v_z$, the kinetic energy is equal to $\frac{1}{2}m(v_x^2 + v_y^2 + v_z^2)$, while $p_x$ is equal to $m v_x$, *etc.*

$$\nabla^2 S + \frac{8\pi^2 m}{h^2}(E - V)S = 0, \tag{11}$$

an equation which we recognize as identical with the *fundamental equation of Schrödinger's theory*. The latter is hence actually obtained from Heisenberg's commutation rule in its application to operators.

If we consider, instead of a single particle, a *system of particles*, and denote by $p_{nx}$, $p_{ny}$, $p_{nz}$, the momentum components of the $n$th particle, eqn. (8) becomes

$$\sum_n \frac{1}{2m}(p_{nx}^2 + p_{ny}^2 + p_{nz}^2) + V - E = 0, \tag{12}$$

$V$ being now the total potential energy and $E$ the total energy of the system. Eqn. (10) then assumes the form

$$\left\{ -\frac{h^2}{8\pi^2} \sum_n \frac{1}{m_n} \nabla_n^2 + V - E \right\} S = 0, \tag{13}$$

and from this we obtain Schrödinger's generalized differential equation, as quoted in § 97 eqn. (26),[6]

$$\sum_n \frac{1}{m_n} \nabla_n^2 S + \frac{8\pi^2}{h^2}(E - V)S = 0. \tag{14}$$

Matrix mechanics exhibits many points of relationship not only with Schrödinger's theory, but also with a third system of atomic dynamics which likewise originated in 1925, *viz.*, the *mechanics of q-numbers*, devised by *Dirac*.[7] This theory is based on the assumption that quantum-mechanical quantities may be represented by a species of "supercomplex" numbers, termed by Dirac $q$-numbers. These differ from ordinary numbers (in Dirac's nomenclature $c$-numbers) by *not obeying the commutative law of multiplication*. Dirac regards as $q$-numbers not only coordinates, momenta, and energy (as the function of coordinates and momenta), but also *time*.[8] The results yielded by

[6] The solutions afforded by Schrödinger's differential equation for any special problem permit of the calculation of the matrix terms in the corresponding quantum-mechanical problem, the necessary transformations being not very difficult. *Cf.* Schrödinger's "*Collected Papers, etc.*," and C. Eckart, *Phys. Rev.*, **28** (1926), p. 711.

[7] P.A.M. Dirac, *Proc. Roy. Soc.*, *A* **109** (1926), p. 642, and **110** (1926), p. 561.

[8] Dirac's theory, by ascribing to time and spatial coordinates the same character, *viz.*, that of $q$-numbers, permits of a much easier approach to the theory of relativity than does matrix mechanics, in which spatial coordinates appear as matrices, whereas time is regarded as an ordinary arithmetical quantity.

Dirac's theory in its application to the special problems of atomic physics agree, in general, with those derived from Schrödinger's and Heisenberg's theories.

### § 99. The Statistical Interpretation of Material Waves.

The analogy between light-corpuscles and electrons suggests a search for an interpretation of *Schrödinger's field-scalar* ($S$) similar to that given, from the point of view of the *theory of light-quanta*, to the quantities characterizing the *optical wave field*. A *corpuscular* theory of light can give only a *statistical explanation* of *intensity* and *wave amplitude*, the square of which determines the intensity. The value of the amplitude ascribed by the wave theory to any point in the field can be viewed by the corpuscular theory only as a *measure of the probability of a coincidence of a light-quantum* with that point. The greater the probability the larger is the amplitude, and hence, the intensity of the light at the point in question.

This conception renders superfluous the unintelligible assumption that light-quanta can destroy one another. It does show, however, that the probability of the coincidence of light-quanta at a point may be so small, in spite of the proximity of sources of light, that the point remains in darkness. A statistical interpretation of this kind of a light-quantum field, moreover, does not infringe on the compatibility of the strict laws governing the observed phenomena of optical interference and diffraction, with the hypothesis that the motion of an *individual light-quantum* may be *undetermined*.

*Born* [1] has now ascribed a similar statistical meaning to Schrödinger's waves, interpreting Schrödinger's *field-scalar* as the *probability of the coincidence of a material particle* with a given point in the field. The accuracy of this assumption has been proved by observations on the photo-electric effect. It follows from wave-mechanics that a body on which a light-wave impinges must, in some circumstances, emit a material wave. In Bohr's view this wave must determine the *distribution of direction* among the *swarm of photo-electrons* emitted by the body. As we shall see in a later chapter, relations, which are well supported by experiment, are actually obtained in this way for the directional distribution.

The statistical interpretation of material waves also

[1] M. Born, *Zeitschr. f. Phys.*, **37** (1926), p. 863, and **38** (1926), p. 803.

removes the apparent contradiction that Schrödinger's theory, which postulates continuity and a space-time representation, is identical with quantum mechanics, which rejects both those concepts. For, although the *elementary* processes of physics are perhaps in no way determined, yet the *probabilities* to be ascribed on *statistical* grounds to these *individual* processes may be treated as *continuously* variable and *determined* quantities. They may satisfy even a differential equation of "classical" type, such as Schrödinger's fundamental equation.[2]

### § 100. The Angular Momentum of Electrons and Atoms.

In 1915 *Sommerfeld* and *W. Wilson*[1] simultaneously, but quite independently, reached the important conclusion that a *second quantum number* must be introduced into atomic theory, if the latter were not to remain restricted to the special case of uniform circular motion. *Any plane, periodic, motion* possesses, as is known, *two degrees of freedom*, so that the description of such a motion requires two mutually independent positional coordinates (in the general sense of this word).[2] If the kinetic energy be represented in terms of these two positional coordinates, we obtain the corresponding *momenta*[3] by the customary partial differentiation with respect to their time-derivatives. According to Sommerfeld's and Wilson's hypothesis, if the two coordinates be denoted by $q_1$ and $q_2$, and the corresponding momenta by $p_1$ and $p_2$, we must have by analogy to Bohr's quantum condition—§ 96 eqn. (3)—

$$(1) \qquad \oint p_1 \, ds = n_1 h \, ; \oint p_2 \, ds = n_2 h.$$

The integrals are to be taken over the orbit, and $n_1$ and $n_2$ denote two different *integral quantum numbers*.

We may suitably use *plane polar coordinates* ($r$ and $\varphi$)

[2] Schrödinger himself assumed that the field-scalar ($S$), introduced by him into physics, represented a measure of the mean electrical density of charge, and that it was just the periodic variations of the mean density of charge that generated radiation in the sense of the classical theory of fields. *Cf.* ch. xvii.

[1] A. Sommerfield, *Ber. der Münchner Akad.*, 1915, p. 425 and 1916, p. 131 : W. Wilson, *Phil. Mag.*, **29** (1915), p. 795, and **31** (1916), p. 156.

[2] *Cf.* Vol. I., § 22.

[3] It is no more necessary that "generalized" momenta should have the dimensions of a product of mass and velocity than that generalized coordinates should have the dimensions of length. The product of a coordinate and the corresponding momentum, however, has always the dimensions of action.

in describing a plane periodic motion, and we have then for the kinetic energy [4]

$$L = \frac{m}{2}(\dot{r}^2 + r^2\dot{\varphi}^2) \tag{2}$$

($\dot{r}$ and $\dot{\varphi}$ denoting the derivatives of $r$ and $\varphi$ with respect to the time, and $m$ the mass). From this relation we find

$$p_r = \frac{\partial L}{\partial \dot{r}} = m\dot{r} \tag{3}$$

and

$$p_\varphi = \frac{\partial L}{\partial \dot{\varphi}} = m r^2 \dot{\varphi} = mf, \tag{4}$$

where $f$ denotes the moment of velocity, by which we mean, as ordinarily, twice the magnitude of the surface described by the radius-vector in a unit of time.[5] According to the well-known law of Kepler,[6] the moment of velocity (and hence $p_\varphi$ also) is *constant* when the force is directed towards a fixed point, and when it is inversely proportional to the square of the distance from this point. Again, the product of the moment of velocity and the mass represents the angular momentum. One of the two quantum numbers ($n_\varphi$) is therefore a measure of the *angular momentum of the periodic orbital motion of the electron*. We shall denote this number by $l$.

Let us consider in particular an atom in which only a *single electron* is travelling around the atomic nucleus. Calculations in atomic mechanics, into which we cannot enter here,[7] lead to the result that, in the state characterized by the two quantum numbers $n_r$ and $n_\varphi$, the energy is given approximately by [8]

[4] We have

$$x = r\cos\varphi, \qquad y = r\sin\varphi,$$

from which follow by differentiation with respect to the time

$$\dot{x} = \dot{r}\cos\varphi - r\sin\varphi \,.\, \dot{\varphi},$$
$$\dot{y} = \dot{r}\sin\varphi + r\cos\varphi \,.\, \dot{\varphi}.$$

On squaring these two equations and adding them together, we obtain the value of the square of the velocity; and multiplication of this by half the mass gives the kinetic energy, as in eqn. (2).

[5] In an element of time $dt$, the radius-vector describes an isosceles triangle the side of which is equal to $r$ and the base to $r\dot{\varphi}\,dt$.

[6] *Cf.* Vol. I., § 8.

[7] *Cf.* A. Sommerfeld's "*Atomic Structure and Spectral Lines*" (London: Methuen, 1925).

[8] If the dependence of mass on velocity in accordance with the theory of relativity be taken into account, even in an atom with a single electron the energy depends in a slight degree on the quantum number $l$.

$$E = -\frac{C}{(n_r + n_\varphi)^2}, \tag{5}$$

where $C$ is a constant characteristic of the atom. The energy is therefore approximately determined by a quantum number $n$, where

$$n = n_r + n_\varphi . \tag{6}$$

When *several electrons* are revolving round the nucleus, circumstances are quite different. The other electrons then generate an *accessory field* which is superimposed upon the "Coulomb" field due to the nucleus. Exact calculation shows that, in consequence of this, each revolving electron executes a *rosette-like motion*,[9] the perihelion of the Kepler ellipse undergoing a precession, whilst the energy of a given state depends, not only on the quantum-number $n$, but also on the number $l$ determining the angular momentum.

Considerations of atomic mechanics, which must be omitted here,[10] lead also to the conclusion that, in the transition of an electron from one quantum state to another in the atom, the number $l$ must alter by unity, *i.e.*, that

$$\Delta l = \pm 1 , \tag{7}$$

whilst the change in value of the other number $n$ is not subject to any such restriction. Observations on optical spectra do, in fact, show that the *spectral terms* of an element can be arranged in *different sets*, and that terms in the same set do not in general combine with one another, combinations occurring only between neighbouring sets. We must thus obviously ascribe to the several sets different values of the number $l$, which determines the angular momentum, whilst within any one set the individual terms are to be distinguished by different values of the number $n$. Observation further shows that, among the sets of terms, the set with the lowest value of $l$ occupies a special position in that its terms, unlike those of the other sets, are never double or multiple, but *always simple*. The theory of multiplicity, to be discussed in the next section, interprets this characteristic by ascribing to this set a value of $l = 0$.

Again, it follows from atomic mechanics (we cannot enter into details here) that any given value of the number $n$ is associated with a total possible number of $n$ different values for the number $l$, these being *successive integers* from

[9] *Cf.* Sommerfeld, *l.c.*
[10] *Cf.* Sommerfeld, *l.c.*

zero upwards. For a given value of $n$, the following values for $l$ are possible :—

$$l = 0, 1, 2 \ldots . (n-1). \tag{8}$$

The quantity $n$ is now generally known as the *principal quantum number*, and the quantity $l$, which determines the *angular momentum*, is called the *subordinate quantum number*.

As has already been stated, the new quantum and wave mechanics rejects the concept of definite electronic orbits. It has, however, also yielded the result (admittedly after some very complicated calculations [11]) that the principal quantum number is to be associated with a subordinate quantum number which may assume any of the values given in eqn. (8). In accordance with the conclusions of the new quantum theory [§ 97 eqn. (25)], the angular momentum, as determined by the subordinate quantum number $l$, is expressed by the equation

$$U(l) = \frac{h}{2\pi}\sqrt{l(l+1)}. \tag{9}$$

The translational motions alone of the electrons in an atom are insufficient to determine its *total angular momentum*, as was shown in 1925 by *Uhlenbeck* and *Goudsmit*.[12] We must also take into account a *proper rotation* of the electrons, or "spin," which imparts to them the qualities of minute *magnets*. Uhlenbeck and Goudsmit associated a *quantum-number one-half* with the spin of an electron, distinguishing the *two possibilities of opposite rotations*, for which the algebraical difference between the quantum numbers should be unity. The existence of an analogous proper rotation for protons may be concluded from the band spectrum of a hydrogen molecule, as will be shown subsequently.

Every electron in an atom may thus be associated with two vectors, one representing the angular momentum of the revolution, and the other the angular momentum of the spin. We shall denote these two vectors by the symbols $\mathbf{l}_i$ and $\mathbf{r}_i$, the subscript $i$ distinguishing between the several electrons. By forming the *vectorial sum* of all these quantities for all the electrons in the atom, we obtain the total angular momentum of the latter. Now, in 1925,

[11] *Cf.* M. Born's "*Problems of Atomic Dynamics*" (Cambridge, U.S.A.; 1926), Part I., lect. 16.

[12] G. E. Uhlenbeck and S. Goudsmit, *Naturwiss.*, **13** (1925), p. 953, and *Nature*, **117** (1926), p. 264.

*Russell* and *Saunders* [13] propounded the fundamental postulate that the individual vectors shall be so *directed* that, in the sense of eqn. (9), their resultant may be represented by a whole or a half quantum number. In other words, the *vectorial frame of the atom* must be so arranged that the *resultant* is itself "*quantized*." Denoting this resultant vector by the symbol $\mathbf{j}$ and the corresponding quantum number by $j$, we must have [14]

$$|\mathbf{j}| = \frac{h}{2\pi}\sqrt{j(j+1)}\,. \tag{10}$$

Furthermore, it follows from spectroscopic experiments that, not only must the quantity $\mathbf{j}$ in itself be "quantized," but also that the same condition must apply to $\mathbf{l}$, the resultant of all the $\mathbf{l}_i$. Thus we must also have

$$|\mathbf{l}| = \frac{h}{2\pi}\sqrt{l(l+1)}\,, \tag{11}$$

with integral values for $l$.[15] Finally we must assume that the vectors $\mathbf{r}_i$ are parallel in either the same or the opposite

[13] H. N. Russell and F. A. Saunders, *Astrophys. Journ.*, **61** (1925), p. 38.

[14] $j$ does *not* represent the magnitude of the vector $\mathbf{j}$.

[15] Introducing, for brevity, the symbol

$$[x] = \sqrt{x(x+1)}\,,$$

we may readily deduce that the *largest integral value or odd multiple of one-half*, which is possible for the *resultant of two vectors* of magnitudes $x$ and $y$, is equal to the expression $(x+y)$. The simplest way of demonstrating this result is to remember that the *error* consequent on an identification of $[x]$ with $x$ becomes *smaller* as $x$ increases until, for infinite $x$, it vanishes asymptotically. If, therefore, the above rule holds for the *smallest* possible values of $x$ and $y$ (other than zero), *i.e.*, for $x = y = \frac{1}{2}$, we may expect it to hold quite generally. Now, for $x = y = \frac{1}{2}$,

$$[x] = [y] = \sqrt{\tfrac{1}{2}\cdot\tfrac{3}{2}} = \tfrac{1}{2}\sqrt{3}\,,$$

so that

$$[x] + [y] = \sqrt{3} = 1{\cdot}732\,.$$

Again, with these values,

$$x + y = \tfrac{1}{2} + \tfrac{1}{2} = 1\,,$$

and hence

$$[x+y] = \sqrt{1{\cdot}2} = 1{\cdot}414\,,$$

while

$$[x+y+\tfrac{1}{2}] = \sqrt{\tfrac{3}{2}\cdot\tfrac{5}{2}} = \tfrac{1}{2}\sqrt{15} = 1{\cdot}935\,.$$

For the *lowest* possible values of $x$ and $y$, *viz.*, for $x = y = \frac{1}{2}$, the largest integral value, or odd multiple of one-half, for the *resultant quantum number* proves to be equal to $(x+y)$, the value $(x+y+\frac{1}{2})$ being even then *too large*. We may thus conclude quite generally that, when two vectors with quantum numbers $r$ and $l$ are compounded, the largest possible resultant quantum number will be $(r+l)$, and similarly the smallest will be $|r-l|$.

sense, so that, with $z$ electrons, we have the following possibilities for the resultant quantum number $r$ :—

$$(12)\quad \begin{cases} r=\frac{1}{2}, \frac{3}{2} \ldots \frac{z}{2} & \text{—odd multiples of one-half for odd values of } z, \\ r=0, 1 \ldots \frac{z}{2} & \text{—integral values for even values of } z. \end{cases}$$

Considerations of atomic mechanics (which must be omitted here) have shown that the quantum number $j$ can *change only by unity or zero ;* so that there are *three possible transitions, viz.,*

$$(13)\qquad j \rightarrow j+1;\ j \rightarrow j;\ j \rightarrow j-1.$$

### § 101. The Multiplicity of Spectral Terms.

It has been found from spectroscopic observations that, in the absence of an external field, a principal quantum number and a subordinate number suffice for the determination of spectral terms, provided that the *fine structure of spectra* be left out of account. By fine structure we mean the phenomenon that many spectral lines, which in a spectroscope of low power appear as simple lines, in an apparatus of high resolution are seen to consist of *groups of closely adjacent lines.* From this phenomenon we must evidently conclude that *spectral terms* in general are *composite,* so that they represent so-called *multiplets.* According to the number of components we speak of *doublets, triplets, quartets, etc.,* whilst simple terms are also called *singlets.* The individual components of a multiplet are hence distinguished only by the value of the quantity $j$, known as the *inner quantum number,* for the principal and the subordinate quantum number $l$ are *the same* for all the components of a term multiplet. The inner quantum number was introduced into atomic theory by *Sommerfeld* in 1920, in order to afford an explanation for fine structure.

The fundamental rules of spectral multiplicity can be readily derived from the concept of the *spin of electrons.* Since the *inner quantum-number* $j$ represents the *vectorial resultant* of the quantum numbers $l$ and $r$, if follows from the discussion in the previous section that the largest value attainable by $j$ is equal to $(l + r)$. The other possible values of $j$ are zero and all positive numbers less than $(l + r)$—the differences being integral—but not smaller than $| l - r |$. The *maximum multiplicity* of the *element* concerned is therefore determined by the number $(2r + 1)$. When $l$ is less than $r$, however, the actual multiplicity is maller, because negative values of $j$ are excluded. The

case in which $l = 0$ occupies a special position, for then, since the vector **l** vanishes, **j** is necessarily *identical* with **r**. Hence, for $l = 0$, only singlet terms are possible, and conversely the experimental fact that, in the spectra of all elements, singlet terms occur justifies our commencing the enumeration of $l$-values with zero. The possible $j$-values obtained according to these rules for given values of $l$ and $r$ are reproduced in Table V for $r = 0$ to $\frac{7}{2}$ and $l = 0$ to 3.

TABLE V

*The Possible Values of $j$ for Given Values of $l$ and $r$*

| | $l = 0$. | $l = 1$. | $l = 2$. | $l = 3$. | Maximum multiplicity for given $r$. |
|---|---|---|---|---|---|
| $r = 0$ | 0 | 1 | 2 | 3 | Singlet. |
| $r = \frac{1}{2}$ | $\frac{1}{2}$ | $\frac{1}{2}\ \frac{3}{2}$ | $\frac{3}{2}\ \frac{5}{2}$ | $\frac{5}{2}\ \frac{7}{2}$ | Doublet. |
| $r = 1$ | 1 | 0 1 2 | 1 2 3 | 2 3 4 | Triplet. |
| $r = \frac{3}{2}$ | $\frac{3}{2}$ | $\frac{1}{2}\ \frac{3}{2}\ \frac{5}{2}$ | $\frac{1}{2}\ \frac{3}{2}\ \frac{5}{2}\ \frac{7}{2}$ | $\frac{3}{2}\ \frac{5}{2}\ \frac{7}{2}\ \frac{9}{2}$ | Quartet. |
| $r = 2$ | 2 | 1 2 3 | 0 1 2 3 4 | 1 2 3 4 5 | Quintet. |
| $r = \frac{5}{2}$ | $\frac{5}{2}$ | $\frac{3}{2}\ \frac{5}{2}\ \frac{7}{2}$ | $\frac{1}{2}\ \frac{3}{2}\ \frac{5}{2}\ \frac{7}{2}\ \frac{9}{2}$ | $\frac{1}{2}\ \frac{3}{2}\ \frac{5}{2}\ \frac{7}{2}\ \frac{9}{2}\ \frac{11}{2}$ | Sextet. |
| $r = 3$ | 3 | 2 3 4 | 1 2 3 4 5 | 0 1 2 3 4 5 6 | Septet. |
| $r = \frac{7}{2}$ | $\frac{7}{2}$ | $\frac{5}{2}\ \frac{7}{2}\ \frac{9}{2}$ | $\frac{3}{2}\ \frac{5}{2}\ \frac{7}{2}\ \frac{9}{2}\ \frac{11}{2}$ | $\frac{1}{2}\ \frac{3}{2}\ \frac{5}{2}\ \frac{7}{2}\ \frac{9}{2}\ \frac{11}{2}\ \frac{13}{2}$ | Octet. |
| Maximum multiplicity for given $l$. | Singlet. | Triplet. | Quintet. | Septet. | |

*Simple numerical relations* have been experimentally obtained for the *frequency intervals* within a multiplet. If $\Delta\nu$ be the interval between two adjacent components with inner quantum numbers $j$ and $j'$ $(= j + 1)$, it is found that $\Delta\nu$ is proportional to $j'$. Hence, if a multiplet be represented by the series of inner quantum numbers

$$j_1, j_2, j_3 \cdots j_s,$$

the $(s - 1)$ intervals in this $s$-fold term bear the same ratios to one another as

(1) $$j_2 : j_3 : j_4 \cdots : j_s.$$

For example, the ratios for the six intervals in the septet $(l = 3, r = 3)$ of Table V are 1 : 2 : 3 : 4 : 5 : 6, whilst those

for the intervals in the septet ($l = 3$, $r = \frac{7}{2}$) are $3 : 5 : 7 : 9 : 11 : 13$.

As Landé has shown,[1] a theoretical basis is afforded to this so-called *Landé's interval rule* (that $\Delta\nu$ is proportional to $j'$) if, as may justifiably be assumed, the interaction between $l$ and $r$ be associated with a potential energy proportional to the *cosine* of the angle included by the vectors $\mathbf{l}$ and $\mathbf{r}$. Since the vectors $\mathbf{j}$, $\mathbf{l}$, and $\mathbf{r}$ form a triangle, $\mathbf{j}$ being the resultant of $\mathbf{l}$ and $\mathbf{r}$, we have

$$|\mathbf{j}|^2 = |\mathbf{l}|^2 + |\mathbf{r}|^2 - 2\,|\mathbf{l}|\,|\mathbf{r}|\cos(\mathbf{l}, \mathbf{r}), \tag{2}$$

or, in accordance with § 97 eqn. (25),

$$\cos(\mathbf{l}, \mathbf{r}) = \frac{l(l+1) + r(r+1) - j(j+1)}{2\sqrt{r\,l\,(r+1)\,(l+1)}}. \tag{3}$$

On the above assumption, the energy corresponding to a component of the multiple term is to be equated to $E$, where

$$E = E_0 + K\cos(\mathbf{l}, \mathbf{r}); \tag{4}$$

both $E_0$ and $K$ must be independent of the inner quantum number. Hence, on forming eqn. (4) for two adjacent components in a multiple term, subtracting, and dividing the difference by the elementary quantum of action, we find

$$|\Delta\nu| = C\,[j'(j'+1) - j(j+1)], \tag{5}$$

where $C$ denotes another quantity independent of $j$; or, since $j' = j + 1$,

$$|\Delta\nu| = 2\,C\,j'; \tag{6}$$

and this actually expresses the proportionality between the frequency interval and the larger of the two inner quantum numbers.

### § 102. The Resolution of Spectral Terms in a Magnetic Field.

The three quantum numbers (principal, subordinate, and inner) previously considered suffice for the description of atomic states only so long as the atom is not subject to any external field. As *Landé* was the first to recognize in 1921, an investigation of the behaviour of atoms in *magnetic fields* demands the introduction of an additional quantum number, which he called the *magnetic quantum number*. From the point of view of classical mechanics, when a system of electrically charged particles is situated in a homogeneous magnetic field, the *component of the total angular momentum coinciding with the direction of the field* must itself be *constant*. The quantization of the total

[1] A. Landé, *Zeitschr. f. Phys.*, **15** (1923), p. 189.

angular momentum ($U$) of the atom in atomic mechanics suggests that we should also regard the constant component in the direction of the field ($U_z$) as *quantized*, and that we should allot to it a special "magnetic" quantum number, which may be denoted by $m$. This number must at any rate be governed by the condition that the magnitude of $U_z$ cannot be greater than that of $U$. Therefore

$$-j \leqq m \leqq +j. \tag{1}$$

Accurate calculations of quantum mechanics show that, just as in the older quantum theory,[1]

$$U_z = m \frac{h}{2\pi}, \tag{2}$$

and that the same *possibilities of transition* exist for the magnetic as for the inner quantum number, *viz.*,

$$m \rightarrow m+1, \quad m \rightarrow m, \quad m \rightarrow m-1. \tag{3}$$

Consequently, $m$ can assume all values from $(-j)$ to $(+j)$ differing from one another by unity, so that

$$m = (-j), (-j+1) \ldots (j-1), j. \tag{4}$$

The total number of different values of $m$ for a given value of $j$ are therefore $(2j+1)$. To these values correspond $(2j+1)$ different orientations of the atom in the magnetic field, so that there is also a *quantization of direction* in the latter.

Now it follows from the classical theory of the electromagnetic field [2] that a body moving with a velocity $v$, and charged with a quantity of electricity $e$, is subjected, in a magnetic field of field-strength $H$, to a *force* which acts at right angles both to the field and to the direction of motion. The magnitude of the force is expressed by the relation

$$K = \frac{e}{c} v H \sin\gamma, \tag{5}$$

where $\gamma$ is the angle included between the direction of motion and that of the field. Again, it is known from classical mechanics [3] that a body, the mechanical behaviour of which is referred to a *rotating coordinate system*, experiences an apparent accessory force, known as a *Coriolis force*. This force acts at right angles to the direction of

[1] It should be noticed that, even in accordance with the new quantum mechanics, $U_z$ is *not* equal to $\frac{h}{2\pi}\sqrt{m(m+1)}$, as might well be expected.

[2] *Cf.* § 71 eqn. (4).

[3] *Vide* § 14 eqn. (2).

motion and to the axis of rotation, and its magnitude is given by the relation

$$K' = 2\mu v w \sin\gamma', \tag{6}$$

where $\mu$ denotes the mass, $w$ the angular velocity of rotation, and $\gamma'$ the angle included between the direction of motion and that of the axis. Comparing the last two equations, we see that the effect of a magnetic field consists simply in the superposition on the motions, which would be executed by the charged body *in the absence* of the field, of a *uniform precessional motion* of the whole system *about the direction of the field*. It follows from eqns. (5) and (6) that the precession occurs with an angular velocity

$$w = \frac{e}{2\mu c} H, \tag{7}$$

or, if we divide by $2\pi$, with a revolution frequency $\omega H$, where the universal constant

$$\omega = \frac{e}{4\pi\mu c}. \tag{8}$$

Classical physics [4] teaches us that the precessional motion *increases* the *energy* of the system by the amount

$$E' = U' w, \tag{9}$$

where $U'$ is the component of the angular momentum of the system in the direction of the field.

On applying this result to the atom, we find from eqn. (2), since $U'$ and $U_z$ have the same signification,

$$E' = m h \omega H. \tag{10}$$

From this formula, it might be expected that a term corresponding to a frequency $\nu$ would be resolved in a magnetic field according to the equation

$$\nu' = \nu + m\omega H.$$

Actually, however, when the field is comparatively weak, resolution takes place according to the formula

$$\nu' = \nu + g m \omega H, \tag{11}$$

where, as Landé discovered in 1923, the value of the so-called *resolving factor* $g$ is given by

$$g = 1 + \frac{j(j+1) + r(r+1) - l(l+1)}{2j(j+1)}. \tag{12}$$

[4] For a comparatively simple proof *vide* the section on the normal Zeeman effect in Sommerfeld's "*Atomic Structure and Spectral Lines*." The precession caused by a magnetic field is frequently called the "*Larmor precession*," after the first investigator of this process.

Landé's resolution formula can easily be derived from the principle of the vectorial composition of quantum numbers, provided that we make an additional assumption, of which, admittedly, the truth has not yet been strictly demonstrated. This hypothesis is to the effect that, for the magnetic accessory energy, the angular momentum of the electronic *spin* has to be taken into account *twice*. That is to say, $U'$ would not be equal to $m\,h/2\,\pi$, but

$$U' = \frac{m\,h}{2\,\pi} + U'', \tag{13}$$

$U''$ denoting the projection of the angular momentum of the spin, first on the direction of the vector $\mathbf{j}$, and then on the direction of the field.[5]

Thus, if we consider the quantum numbers and replace, in the sense of the newer quantum mechanics, $j$, $r$, $l$, by $j^*$, $r^*$, $l^*$, where

$$j^* = \sqrt{j\,(j+1)}\,, \tag{14}$$

*etc.* [*cf.* § 101 eqn. (3)], we must substitute for $m$ a number $m'$ determined by the relation

$$m' = m + r^* \cos(\mathbf{r}, \mathbf{j}) \cos(\mathbf{j}, \mathbf{H}). \tag{15}$$

Now

$$\cos(\mathbf{j}, \mathbf{H}) = \frac{m}{j^*}\,, \tag{16}$$

and, because the vector $\mathbf{j}$ is the geometrical sum of the vectors $\mathbf{r}$ and $\mathbf{l}$,

$$l^{*2} = r^{*2} + j^{*2} - 2\,r^*\,j^* \cos(\mathbf{r}, \mathbf{j})\,. \tag{17}$$

Hence

$$m' = m\left[1 + \frac{j\,(j+1) + r\,(r+1) - l\,(l+1)}{2\,j\,(j+1)}\right], \tag{18}$$

so that, on putting

$$m' = g\,m, \tag{19}$$

we obtain in fact eqn. (12).

The validity of Landé's resolution formula is found to be limited to *weak magnetic fields*, or, more accurately, to such fields as produce variations in frequency which are small as

[5] The angular momentum of the spin, represented by the quantum number $r$, has a double effect, inasmuch as the quantum number $j$, which corresponds to the total angular momentum (with the component $U'$), is actually the *resultant* of $l$ and $r$.

compared with the frequency intervals in the multiplet in the absence of a field. Stronger magnetic fields *pull* the vectors $\mathbf{r}$ and $\mathbf{l}$ *apart*, so that finally they become *independent of each other*. Each of the two vectors then experiences a quantization of direction, and possesses a *quantized component in the direction of the field:* the two components may be denoted by $m_r$ and $m_l$. If we assume again that, in calculating the magnetic accessory energy, the electronic spin has to be taken *twice* into account, we obtain for the accessory energy the value

$$E = h\,\omega\,(m_l + 2\,m_r)\,H. \tag{20}$$

Now $l$, and hence also $m_l$ is integral; and, since $2\,r$ is always integral, the same must hold for $2\,m_r$. In a *strong* magnetic field, therefore, the frequency $\nu$ characteristic of a spectral line is resolved into a series expressed by

$$\nu' = \nu + m\,\omega\,H, \tag{21}$$

where $m$ is an *integer*.

### § 103. Pauli's Principle.

In 1925 *Pauli* established a principle on the methods of the four-fold quantization of atomic states which has not only proved extraordinarily fruitful in the theory of spectra, but has also provided a key for the interpretation of the periodic system of the elements. *Pauli's principle* states that, in an atom, *no two electrons can have identical values for all four quantum numbers.*

We shall suppose the electrons comprised in the planetary system of an atom to be arranged in *groups* according to the value of the *principal quantum number*, and each group to be divided into *sub-groups* having the same value of the *subordinate quantum number*. Electrons belonging to the same sub-group in a group, and therefore having the same values of principal and subordinate quantum numbers, will be termed *equivalent electrons*.

According to Pauli's principle, the number of equivalent electrons is limited by the restriction that no two of them can agree in the values of the other two quantum numbers which are necessary for four-fold quantization. We may suitably choose for the two other quantum numbers those which correspond to the components of $\mathbf{l}_i$ and $\mathbf{r}_i$ in the direction of a strong magnetic field.[1] On account of the

[1] Exactly the same results are reached if $j$ and $m$ be chosen for the other two quantum numbers.

equivalence, the value of $l_i$ must obviously be the same for all the electrons : we shall denote it by $l$. For the quantum numbers of the components we may write $m_i\,(l)$ and $m_i\,(r)$, and it is to be remembered that the rotating electrons, as elementary magnets, must arrange themselves parallel to the direction of the field in either the same or the opposite sense, so that $m_i\,(r)$ can have only the values $(+\frac{1}{2})$ and $(-\frac{1}{2})$. On the other hand, $m_i\,(l)$ may assume any integral value from $(-l)$ to $(+l)$. There are thus $(2l+1)$ possible values for $m_i\,(l)$, and any one of these may be combined with the value $(+\frac{1}{2})$ or $(-\frac{1}{2})$ of $m_i\,(r)$. For a given value of the subordinate quantum number $l$, the *maximum number of equivalent electrons* is thus found to be

$$N_{\max.}\,(l) = 4l + 2. \tag{1}$$

When the *full number* of electrons are present, $m_i\,(l)$ assumes *every* integral value from $(-l)$ to $(+l)$ twice over, so that

$$\sum m_i\,(l) = 0. \tag{2}$$

In the same circumstances, the values $(+\frac{1}{2})$ and $(-\frac{1}{2})$ occur equally often for $m_i\,(r)$, so that

$$\sum m_i\,(r) = 0. \tag{3}$$

Denoting, now, by $L$ the resultant quantum number of all the vectors $\mathbf{l}_i$, *all* integral values from $(-L)$ to $(+L)$ must be possible for $\Sigma\, m_i\,(l)$. Hence, by eqn. (2),

$$L = 0, \tag{4}$$

and similarly, if $r$ denotes the resultant quantum number of all the vectors $\mathbf{r}_i$, it follows from eqn. (3) that

$$r = 0. \tag{5}$$

Eqns. (4) and (5) show that a *complete sub-group* contributes *nothing* to the vectorial frame of an atom, and this fact evidently justifies the limitation of the theory of spectra to the consideration of those electrons which, in the atom, lie *outside* the completed groups and sub-groups.

For a given *principal quantum number* $n$, all values of $l$ from 0 to $(n-1)$ are possible. The maximum number of electrons in an $n$-quantum group is therefore obtained from the relation

$$N_{\max.}\,(n) = 4\sum_{l=0}^{l=n-1} l + 2n. \tag{6}$$

But

$$\sum_{l=0}^{l=n-1} l = \left(\frac{n-1}{2}\right) n. \tag{7}$$

Hence

$$N_{\text{max.}}(n) = 2\,n^2. \tag{8}$$

A one-quantum group is therefore *completely occupied* with 2, a two-quantum group with 8, a three-quantum with 18, and a four-quantum group with 32 electrons.

## Chapter XIII

# ATOMIC SPECTRA

### § 104. The Spectrum of Hydrogen.

As long ago as 1885, *Balmer* made the important discovery that very *simple numerical relations* exist between the *wave-lengths* of the four *spectral lines* which characterize *hydrogen* in the optical region. The frequencies of these four lines (the red $H_\alpha$, the blue $H_\beta$, and the violet $H_\gamma$ and $H_\delta$) can, in fact, be represented by the formula

$$\frac{\nu}{c} = R\left(\frac{1}{4} - \frac{1}{n^2}\right), \tag{1}$$

where $R$ denotes the number 109,678, and $n$ is given the values 3, 4, 5, or 6.[1] It has been found that Balmer's formula is also obeyed with the greatest exactness for the numerous lines which mark the *continuation into the ultra-violet* of the optical series of hydrogen. Besides this series, of which the limit lies at 3,647 A. and some fifty lines have been measured,[2] *three other hydrogen series* are now known. In the far *ultra-violet* lies a series discovered in 1914 by *Lyman*[3]: this satisfies the formula

$$\frac{\nu}{c} = R\left(1 - \frac{1}{n^2}\right). \tag{2}$$

In the *infra-red* a prediction by *Ritz* was verified by *Paschen's* discovery, in 1909, of the series

[1] Balmer himself wrote his formula as

$$\lambda = C\frac{n^2}{n^2 - 4},$$

where $C$ is a constant equal to $4/R$.

[2] It has, admittedly, been possible to observe this large number of lines only in the *absorption spectrum*. In the emission spectra of vacuum tubes, at most twenty lines have been detected, but an additional thirteen emission lines have been found in the photographs of nebulous stars.

[3] At present the first three lines of this series are known. The line for which $n = 2$ (thus with a wave-number equal to $\frac{3}{4}R$) appears very sharply in all photographic exposures made in the extreme ultra-violet: its wave-length is 1215·7 Å.

$$\frac{\nu}{c} = R\left(\frac{1}{9} - \frac{1}{n^2}\right). \tag{3}$$

Finally, in 1922, *Brackett* discovered in the far infra-red the first two lines of a series corresponding to the formula [4]

$$\frac{\nu}{c} = R\left(\frac{1}{16} - \frac{1}{n^2}\right). \tag{4}$$

The exact way in which the four observed hydrogen series satisfy the above formulæ clearly indicates that the *hydrogen terms* are *inversely proportional to the squares of whole numbers*, so that they may be represented in the form

$$T = \frac{R}{n^2}. \tag{5}$$

In point of fact, the wave-numbers of all the observed lines of hydrogen can be written as the differences between two such terms. It has been possible to determine the constant $R$ from spectroscopic measurements with an accuracy of one in a million, and the value thus found from the hydrogen spectrum (which fact may be expressed by adding a subscript $H$) is [5]

$$R_H = 109{,}677{\cdot}70\ (\pm\ 0{\cdot}04)\ \text{cm}^{-1}. \tag{6}$$

The equivalent values of wave-length, potential, and energy are

$$R : 911{\cdot}8\ \text{Å}. : 13{\cdot}52\ \text{volts} : 2{\cdot}15 \times 10^{-11}\ \text{ergs}. \tag{7}$$

The quantity $R$ is known as *Rydberg's constant*, because, in 1895, Rydberg discovered that this constant has a real significance, not only for the hydrogen spectrum, but also for the spectra of other elements. In 1910, the author of this book was able to show that this fundamental constant of spectroscopy was closely related to the elementary quantum of electricity, the mass of an electron, and the elementary quantum of action.[6] This relation was derived more accurately by *Bohr* in 1913, on the basis of Rutherford's atomic model and of the principle previously discussed, according to which, in a circular motion of an electron, the *angular momentum* multiplied by $2\pi$ is equal to an *integral multiple of the elementary quantum of action*.

[4] On the same occasion, Brackett also discovered the third, fourth and fifth lines of the Paschen series, of which only the first two lines had been previously known.

[5] The value here quoted, which is probably the most accurate available, is based on data of W. V. Houston, *Astrophys. Journ.*, 64 (1926), p. 81.

[6] A. Haas, *Sitz. Ber. der Wiener Akad., Abt.* IIa, 1910, p. 119; also *Physikal. Zeitschr.*, **11** (1910), p. 537.

In order to derive the connection between Rydberg's constant and the fundamental quantities of the electron theory, we shall consider a nucleus, with a positive charge of $z$ elementary quanta, about which a *single electron* is describing a *circular orbit* (we generalize, rather than put $z = 1$ at once, for reasons which will appear in the next section). The forces acting on the revolving electron are, on the one hand, the *electrical attraction of the nucleus*, and, on the other hand, the *centrifugal force*. In accordance with Newton's Third Law of Motion, the *two forces* must be *equal*. By *Coulomb's Law*, the force of attraction is equal to $z\,e^2/a^2$; the centrifugal force is $m\,v^2/a$, where $m$ is the mass of the electron. Equating the two expressions, we find

$$m\,v^2\,a = z\,e^2. \tag{8}$$

Again, since the angular momentum multiplied by $2\pi$ must be an integral multiple of the elementary quantum of action, the following relation must be obeyed :—

$$m\,a\,v = \frac{n\,h}{2\,\pi}. \tag{9}$$

If the *quantum number* ($n$) be given, eqns. (8) and (9) contain as unknowns only the orbital radius ($a$) and the linear velocity ($v$), so that these two magnitudes can be determined from the two equations. We may distinguish by the subscript $n$ the values assumed by these magnitudes for the $n$-quantum state. The value of $a_n$ is obtained by squaring eqn. (9) and dividing by eqn. (8) : we have then

$$a_n = \frac{n^2\,h^2}{4\,\pi^2\,z\,e^2\,m}. \tag{10}$$

The *radii* of the orbits derived from the quantum theory therefore vary as the *squares of the quantum numbers*. Again, on dividing eqn. (9) by eqn. (10), we find

$$v_n = \frac{2\,\pi\,z\,e^2}{n\,h}. \tag{11}$$

The *energy* of the electron is made up of kinetic and potential energy. The *kinetic* energy is found by multiplying half the square of the velocity by the mass ($m$). The *potential* energy follows from Coulomb's Law :—

$$E_{\text{pot}} = -\frac{z\,e^2}{a}. \tag{12}$$

A *negative sign* has to be prefixed because, on account of the attraction between nucleus and electron, an *increase in the*

*distance a* requires an *addition of energy*, and the energy can only increase with $a$ (to which it is inversely proportional) if it be *opposite in sign*. In spite of the negative sign, the energy itself is, of course, not negative, for in eqn. (12) there has still to be added a positive constant which has the same value for all states. Besides, the potential energy represents the *work* which would be required to remove the electron, were it at rest, to an infinite distance.

On comparing eqns. (8) and (12) we have

$$E_{\text{pot}} = -m v^2. \tag{13}$$

The potential energy is therefore twice as large as the kinetic energy, and opposite in sign; hence the *total energy* (the sum of kinetic and potential) is given by

$$E_n = -\frac{m v_n^2}{2}, \tag{14}$$

or, according to eqn. (11), by

$$E_n = -\frac{2\pi^2 z^2 e^4 m}{n^2 h^2}. \tag{15}$$

The *energy values* belonging to the different quantum orbits thus vary *inversely as the squares of the quantum numbers.*

For the *spectral frequency* accompanying a transition between two states we obtain thus the relation

$$\nu = \frac{2\pi^2 z^2 e^4 m}{h^3}\left(\frac{1}{s^2} - \frac{1}{n^2}\right). \tag{16}$$

This equation corresponds to Balmer's generalized formula

$$\frac{\nu}{c} = R\left(\frac{1}{s^2} - \frac{1}{n^2}\right); \tag{17}$$

it is only necessary for the special case $z = 1$ to put

$$R c = \frac{2\pi^2 e^4 m}{h^3}. \tag{18}$$

This formula actually proves to be accurately obeyed when we insert on the right-hand side the values found for $e$, $m$, and $h$ from observations based on different principles. It expresses therefore the relation between Rydberg's constant and the fundamental quantities of atomic theory.[7]

[7] This relation is to be found as eqn. (67) in the author's paper, quoted in footnote (6). The occurrence there of a different numerical factor is due to the employment of Thomson's atomic model, which was at that time in use, instead of Rutherford's model, first devised in 1911.

From the point of view of Bohr's theory, the Balmer series in the emission spectrum is characterized by a two-quantum end state, the red line $H_\alpha$, for example, arising on a transition from a three-quantum to a two-quantum state, the blue line $H_\beta$ on a transition from a four-quantum to a two-quantum state, *etc.* The ultra-violet Lyman series arises, on emission, from a transition to the normal state, the Paschen series from a transition to a three-quantum, and the Brackett series from a transition to a four-quantum end state.

If we give $n$ and $z$ the value unity in the previous equations, we obtain the constants characteristic of the *normal state of the hydrogen atom*. We find from eqn. (11)

$$v^* = \frac{2\pi e^2}{h} = 2{\cdot}188 \times 10^8 \text{ cm. sec.}^{-1}. \tag{19}$$

For the ratio of this velocity to that of light, which represents an important constant, we obtain the value

$$\alpha = \frac{v^*}{c} = 7{\cdot}29 \times 10^{-3}. \tag{20}$$

We shall denote the *radius of the orbit* in the normal state by $a^*$. Now the total energy, which [by eqns. (13) and (14)] represents half the potential energy, is equal to $(-e^2/2a)$, and thus

$$a^* = -\frac{e^2}{2E^*} = -\frac{e^2}{2hRc}. \tag{21}$$

Hence it follows from eqn. (7) that

$$a^* = 5{\cdot}30 \times 10^{-9} \text{ cm}. \tag{22}$$

Finally, we obtain the *revolution frequency* in the normal state on dividing the velocity $v^*$ by the circumference of the orbit. We have

$$\omega^* = \frac{v^*}{2\pi a^*}, \tag{23}$$

or, by eqns. (19) and (21),

$$\omega^* = 2Rc\,; \tag{24}$$

whence we find

$$\omega^* = 6{\cdot}58 \times 10^{15} \text{ sec.}^{-1}. \tag{25}$$

The values here given are in no way to be regarded as exact values of atomic physics in the same sense as are the elemen-

tary quantum of electricity, the mass of an electron, the elementary quantum of action, or the Loschmidt number. For it must not be overlooked that they refer, not to the hydrogen atom as such, but only a certain, especially simple, *model* of that atom.

In the previous considerations, we have determined the states of the hydrogen atom with the help of only a single quantum number, the principal quantum number. The effect of the *subordinate quantum number* $l$ is seen when, as *Sommerfeld* demonstrated in 1915, the *dependence of mass on velocity* in accordance with the theory of relativity is taken into account, and when *doubly quantized, elliptical, orbits* are assumed. Theoretical considerations, which must be omitted here, led Sommerfeld to the conclusion that, through the relativity correction, an $n$-quantum term undergoes an $n$-fold resolution.[8]

Whereas, therefore, the one-quantum fundamental term remains a simple one, according to Sommerfeld's theory the two-quantum term becomes a *doublet*, its *frequency difference* being given by the equation

$$\frac{\Delta\nu_H}{c} = R\,\frac{\alpha^2}{16}. \tag{26}$$

$\alpha$ has the value defined by eqn. (20), so that

$$\frac{\Delta\nu_H}{c} = 0{\cdot}365\ \mathrm{cm}^{-1}. \tag{27}$$

Sommerfeld found for the frequency differences in the triplet of the three-quantum term values equal to 8/81 and 8/27 of $\Delta\nu_H$, and for the higher quantum terms he derived still smaller fractions.

In the visible zone of the spectrum, the frequency difference $\Delta\nu_H$ corresponds to a wave-length difference of about 0·1 Å. It is thus apparent that, for the *spectral lines* themselves, even with high resolving power, only the duplicity of the two-quantum term will be observable, whereas the multiplicity of higher terms will escape observation. Actually, with apparatus of sufficient resolution, the lines of the Balmer series appear as *doublets*, and the frequency difference is found to agree approximately with the value calculated by Sommerfeld.

From the point of view of Uhlenbeck and Goudsmit's hypothesis, it seems remarkable that the two-quantum

[8] *Cf.* Sommerfeld, *l.c.*

term should be only doubly resolved, and not, as might be expected, triply in accordance with the scheme

$$\begin{cases} n = 2,\ l = 0,\ j = \frac{1}{2}; \\ n = 2,\ l = 1,\ j = \frac{1}{2}; \\ n = 2,\ l = 1,\ j = \frac{3}{2}. \end{cases}$$

*Uhlenbeck* and *Goudsmit*,[9] and also *Slater*[10] endeavoured to explain this apparent contradiction by assuming that terms with the same values of $n$ and $j$ *coincided*, even when their $l$-values differed. Thus, of the triplet represented above, only two separate terms would appear.

In point of fact, *Heisenberg* and *Jordan*[11] were able to demonstrate, on the basis of the new *quantum mechanics*, that the *doublet separation* of the terms ($n = 2$, $l = 1$, $j = \frac{1}{2}$ or $\frac{3}{2}$) is identical with the frequency difference, due to the relativity correction, for the distance between the terms ($n = 2$, $l = 0$, $j = \frac{1}{2}$) and ($n = 2$, $l = 1$, $j = \frac{3}{2}$). In other words, the two terms with the inner quantum number $j = \frac{1}{2}$ coincided.

### § 105. The Spectrum of Ionized Helium.

Whereas the charge on the hydrogen nucleus is a single positive elementary quantum of electricity, *α-particles* (as already mentioned) represent positive structures with twice as large a charge and four times as large a mass. Since, on the other hand, the atomic weight of the second lightest element, helium, is four, the idea suggests itself of identifying the α-particle with the *helium nucleus*. A direct confirmation of this idea is found in the fact, which was first established in 1903 by *Ramsay* and *Soddy*, that helium is produced by radioactive substances emitting α-rays.[1]

Therefore, if we put the nuclear charge number $z$ equal to two in Bohr's spectral formula [§104 eqn. (16)] the formula thus specialized must represent the *spectrum of the ionized helium atom*, in which only a single electron revolves round the doubly charged nucleus, the second electron, which is present in the neutral condition, having been removed. For a transition from an $n$-quantum to an $s$-quantum orbit, the frequency produced in ionized helium must be four times as large as that produced in hydrogen, seeing that $\nu$ is

[9] G. E. Uhlenbeck and S. Goudsmit, *Physica*, **5** (1925), p. 266.
[10] J. C. Slater, *Proc. Nat. Acad. Sc. U.S.A.*, **11** (1925), p. 732.
[11] W. Heisenberg and P. Jordan, *Zeitschr. f. Phys.*, **37** (1926), p. 263.
[1] Vide § 124.

proportional to $z^2$. The spectrum of ionized helium would therefore be represented by the formula

$$\frac{\nu}{c} = 4R\left(\frac{1}{s^2} - \frac{1}{n^2}\right), \tag{1}$$

which may also be written

$$\frac{\nu}{c} = R\left[\frac{1}{\left(\frac{s}{2}\right)^2} - \frac{1}{\left(\frac{n}{2}\right)^2}\right]. \tag{2}$$

In 1897, *Pickering* actually discovered in the spectrum of the star $\zeta$ *Puppis* a series corresponding to eqn. (2) with $s = 4$; moreover, in 1912, *Fowler* discovered in a Geissler tube spectrum a series described by eqn. (2) when $s$ is put equal to 3.[2] Originally, both the Pickering and the Fowler series were *erroneously ascribed to hydrogen*. Bohr's theory indicates that they are *helium* series, and it was a triumph for that theory when, in 1914, *Paschen* and *Bartels* succeeded in observing the Fowler series in tubes filled with pure helium and completely uncontaminated by any traces of hydrogen.

According to the formulæ derived for the line spectra of hydrogen and helium, every second line of the Pickering series should really coincide with a line of the Balmer series, for the Pickering series is obtained by putting $s = 4$ in eqn. (2) so that, for even values of $n$, the equation becomes identical with § 104 eqn. (1). Actually, however, the corresponding lines of the Pickering series are found to be slightly displaced in the direction of increasing frequency, and thus towards the violet in the visible region of the spectrum. Although this so-called *displacement towards the violet of the helium lines* might at first appear as a defect in Bohr's theory, it really proved to be a brilliant confirmation of it, when Bohr perfected his original theory by taking the motion of the atomic nucleus into account. For, in reality, the electron does not revolve round the nucleus, but both revolve *about their common centre of gravity*.

Denoting by $x$ and $y$ the respective distances separating the electron and the nucleus from the centre of gravity and by $M$ the mass of the nucleus, we have the ratio

$$x : y = M : m, \quad (x + y = a); \tag{3}$$

whence it follows that

[2] Formerly, two Fowler series were distinguished, according as $n$ is odd or even. This distinction, however, appears entirely superfluous from the point of view of Bohr's theory.

$$x = a\,\frac{M}{M+m}, \quad y = a\,\frac{m}{M+m}. \tag{4}$$

Let the velocity of the electron be $v'$, and that of the nucleus $v''$. Then, since both electron and nucleus have the same angular velocity, which may be denoted by $w$,

$$v' = x\,w, \quad v'' = y\,w. \tag{5}$$

From this we find the value of the kinetic energy to be given by

$$E_{\text{kin}} = \frac{w^2}{2}\,(m\,x^2 + M\,y^2),$$

or, by eqn. (4),

$$E_{\text{kin}} = \frac{a^2\,w^2\,M\,m\,(M+m)}{2\,(M+m)^2}.$$

In order to simplify the subsequent calculations, let us put

$$\frac{M\,m}{M+m} = \mu\,; \tag{6}$$

then

$$E_{\text{kin}} = \frac{a^2\,w^2}{2}\,\mu. \tag{7}$$

Again, the centrifugal force must be equal to the force of attraction, so that

$$m\,x\,w^2 = \frac{z\,e^2}{a^2}. \tag{8}$$

Now since, according to eqns. (4) and (6), $m\,x$ is equal to $a\,\mu$, we have

$$\mu\,a^2\,w^2 = \frac{z\,e^2}{a}, \tag{9}$$

and hence, by eqn. (7),

$$E_{\text{kin}} = \frac{z\,e^2}{2\,a}. \tag{10}$$

The potential energy is given by

$$E_{\text{pot}} = -\,\frac{z\,e^2}{a}, \tag{11}$$

and hence the total energy in the $n$-quantum orbit is

$$E_n = -\,\frac{z\,e^2}{2\,a_n} \tag{12}$$

(in agreement with § 104).

For the *angular momentum* we find the value

$$U = m\,x\,v' + M\,y\,v'' = w\,(m\,x^2 + M\,y^2), \tag{13}$$

so that

$$U = \frac{2\,E_{\text{kin}}}{w},$$

or, by eqn. (7),

$$U = a^2\,w\,\mu. \tag{14}$$

According to the fundamental quantum relation, the angular momentum multiplied by $2\pi$ must be equal to the product of the quantum number and the elementary quantum of action. We have therefore

$$2\,\pi\,a^2\,w\,\mu = n\,h. \tag{15}$$

On squaring this equation and dividing by eqn. (9), we obtain

$$a_n = \frac{n^2\,h^2}{4\,\pi^2\,z\,e^2\,\mu}. \tag{16}$$

Hence, in accordance with eqn. (12), the energy is given by

$$E_n = -\frac{2\,\pi^2 z^2 e^4\,\mu}{n^2\,h^2}. \tag{17}$$

Comparing this formula with § 104 eqn. (15), we recognize a complete agreement, except that $m$ is now replaced by the quantity $\mu$. If we define the constant $R$ by § 104 eqn. (18), *Bohr's improved spectral formula* becomes

$$\frac{\nu}{c} = z^2\,R\,\frac{\mu}{m}\left(\frac{1}{s^2} - \frac{1}{n^2}\right),$$

or,

$$\frac{\nu}{c} = \frac{R}{1 + \frac{m}{M}}\left[\frac{1}{\left(\frac{s}{z}\right)^2} - \frac{1}{\left(\frac{n}{z}\right)^2}\right]. \tag{18}$$

Thus, neither in the case of hydrogen nor in that of helium can the constant of the line spectrum be exactly equal to the *Rydberg constant*, as defined by § 104 eqn. (18), in the narrower sense of the term. Rydberg's constant may be denoted by $R_\infty$, because we obtain it when the atomic nucleus is assumed to be infinitely heavy in comparison with the revolving electron.

We shall denote the constants of the hydrogen and the helium spectra by $R_H$ and $R_{He}$, and the nuclear masses of the

two atoms by $M_H$ and $M_{He}$ respectively. By eqn. (18) we have then the relations

$$(19) \qquad \begin{cases} R_H = \dfrac{R_\infty}{1 + \dfrac{m}{M_H}}, \\ R_{He} = \dfrac{R_\infty}{1 + \dfrac{m}{M_{He}}}. \end{cases}$$

But the mass of the helium nucleus is approximately four times that of the hydrogen nucleus, so that

$$(20) \qquad R_{He} > R_H ;$$

and the explanation of the *displacement towards the violet of the helium lines* is at once evident.

Exact measurements (quoted in part in § 104) have given as values for the two quantities $R_H$ and $R_{He}$

$$(21) \qquad \begin{cases} R_H = 109{,}677{\cdot}70\ (\pm\ 0{\cdot}04), \\ R_{He} = 109{,}722{\cdot}14\ (\pm\ 0{\cdot}04). \end{cases}$$

Again, the *atomic weights* of hydrogen and helium are

$$(22) \qquad \mathrm{H} = 1{\cdot}0077,\ \mathrm{He} = 4{\cdot}00.$$

We may therefore put

$$(23) \qquad \frac{M_{He}}{M_H} = 3{\cdot}97,$$

although, in a more exact calculation, we should also have to take into account the amount contributed toward the total atomic weight by the electron which revolves around the nucleus. Let us put, for brevity,

$$(24) \qquad \frac{m}{M_H} = x.$$

Then, by eqn. (19),

$$(25) \qquad \frac{R_H}{R_{He}} = \frac{1 + \dfrac{x}{3{\cdot}97}}{1 + x}.$$

From this it follows that

$$(26) \qquad x = \frac{R_{He} - R_H}{R_H - \dfrac{1}{3{\cdot}97} R_{He}},$$

and hence

$$(27) \qquad \frac{1}{x} = \frac{M_H}{m} = \frac{109{,}678 - 27{,}638}{44{\cdot}44} = 1846.$$

From the *displacement towards the violet of the helium lines*, Bohr's theory thus enables us to *calculate* with great accuracy the *ratio between the masses of the proton and of the electron.*

Eqn. (27) may also be written in the form

$$1846 = (e/m) : (e/M_H). \tag{28}$$

Now the quotient $e/M_H$ is known from electrolytic measurements (*cf.* § 87), and hence *from the displacement toward the violet of the helium lines* we can also *calculate the specific charge of the electron.* The value thus obtained ($5{\cdot}30 \times 10^{17}$ E.S.U./gram) agrees exactly with that derived from direct measurements on the electric and the magnetic deflection of cathode rays.

On introducing the value given by eqn. (27) into the first of eqns. (19), we find

$$R_\infty = R_H \times 1{\cdot}000542 = 109{,}737{\cdot}1\,; \tag{29}$$

and hence, by § 104 eqn. (18),

$$109{,}737{\cdot}1 \times c = \frac{2\pi^2 e^4 m}{h^3}. \tag{30}$$

Eqns. (28) and (30) attain a still higher degree of accuracy when we also take into account the share contributed towards the atomic mass by the revolving electron. In that case, eqns. (26) and (28) are replaced by more complicated formulæ. In accordance with §§ 86 and 87, we can regard the values of $e$ and of $M_H$ as given with very great accuracy by Millikan's determinations and electrolytic measurements. If we then solve the improved eqns. (28) and (30) with respect to $m$ and $h$, we obtain for these two quantities also more accurate values than would be possible from measurements on cathode rays, or on the photo-electric effect. As *Flamm* has shown,[3] we find

$$h = 6{\cdot}545 \times 10^{-27}\ \text{erg. sec.}\ (\pm\, 0{\cdot}012 \times 10^{-27}), \tag{31}$$

$$\frac{e}{m} = 5{\cdot}304 \times 10^{17}\ \text{E.S.U./gram.}\ (\pm\, 0{\cdot}009 \times 10^{17}), \tag{32}$$

and finally,

$$m = 9{\cdot}00 \times 10^{-28}\ \text{gram.}\ (\pm\, 0{\cdot}02 \times 10^{-28}). \tag{33}$$

With respect to its *fine structure*, the spectrum of ionized helium may be expected to exhibit analogous relations to those obtaining for the hydrogen spectrum. Both Sommerfeld's theory, which is based on the mass formula of relativity, and the concept of rotating electrons lead to the con-

[3] Flamm, *Physikal. Zeitschr.*, **18** (1917), p. 515.

clusion that the *resolution is proportional to the fourth power of the nuclear charge number*, so that, for *helium*, it must be *sixteen times as large as for hydrogen*. The spectrum of ionized helium thus offers the possibility of indirectly studying the fine structure of the hydrogen spectrum with much greater accuracy than is attainable by direct observation. In point of fact, calculations, based on the helium spectrum, of the frequency difference of the hydrogen doublet have yielded the theoretically derived value—§ 104 eqn. (27)—and have also not come into conflict with Slater's hypothesis, quoted in the previous section.

### § 106. The Spectra of Atoms with One Valency Electron.

A few years after Balmer had established his fundamental formula for the hydrogen spectrum, *Kayser* and *Runge* discovered simple *series relations* also in the *spectra of the alkali metals*. *Rydberg* succeeded in expressing these relations quantitatively by formulæ which exhibited a strong resemblance to Balmer's formula, and in which a fundamental part was played by the so-called *Rydberg constant*, already known from the hydrogen spectrum. Series relations similar to those for the alkali metals were soon established for the elements in the second and third groups of the periodic system.

The common mark of distinction of all optical *series* is (as has already been mentioned) the *decreasing intensity* with *decreasing wave-length*, together with the *convergence of the lines*, also with decreasing definition, towards a *limit* corresponding to the final level in the case of emission. In those elements for which an arrangement of the optical spectral lines into series was possible, the most important series were soon recognized to be three in number, and these were called the *principal series*, the *first* or *diffuse subordinate series*, and the *second* or *sharp subordinate series*. The two subordinate series, which are clearly distinguished by the different sharpness of their lines, are found to have a *common limit* which, for the elements in the first two groups of the periodic system, always has a greater wave-length than the limit of the principal series. Subsequently, a fourth series of comparatively large wave-lengths was added to the three above-named series, and this, by an ill-chosen expression, is known as the *fundamental series* or, after one of its discoverers, as the *Bergmann series*.[1]

[1] The real discoverer of the "Bergmann series" was *Saunders*.

Later considerations have shown that the frequencies of the most important lines in the series named can be represented as differences between the members of *four series of terms;* these are distinguished by the initial letters of the words "sharp," "principal," "diffuse," and "fundamental" as *s*-, *p*-, *d*-, and *f*-terms. The above four series arise on emission from the transitions represented in the following scheme :—

$$(1)\qquad \begin{cases} \text{Principal series} \quad . \quad . \quad . & p \rightarrow s, \\ \text{1st (diffuse) subordinate series} \; . & d \rightarrow p, \\ \text{2nd (sharp) subordinate series} \; . & s \rightarrow p, \\ \text{Fundamental series} \quad . \quad . & f \rightarrow d. \end{cases}$$

Since the *s*-terms prove to be always simple, it follows from the discussion in a previous section (§ 101) that we must evidently ascribe the subordinate quantum number $l = 0$ to them. Again, since (as already mentioned) $l$ must change by unity on a quantum transition, and the *s*-terms normally combine only with *p*-terms, the value $l = 1$ must be associated with the latter. Furthermore, *p*-terms combine not only with *s*-terms, but also with *d*-terms, and the latter combine both with *p*-terms and *f*-terms. Consequently, the values of the *subordinate quantum number* for the optical spectral terms must be

$$(2)\qquad \begin{cases} l = 0 & \text{for an } s\text{-term}, \\ l = 1 & \text{for a } p\text{-term}, \\ l = 2 & \text{for a } d\text{-term}, \\ l = 3 & \text{for an } f\text{-term}. \end{cases}$$

The terms for which $l$ is to be put $= 4$ and 5 are known as *g*- and *h*-terms, and so on.[2]

Observation shows that, the *lower* the *subordinate quantum number* of the *initial state* on emission, the *sharper* is the spectral line produced. The lines in the second subordinate series are therefore the sharpest, those in the principal series being less sharp; in the first (or diffuse) subordinate series the lines are, as the name indicates, ill defined, and in the fundamental series they are least sharp of all.

It has been found that the individual terms of the alkali metals may be represented by the formula

$$(3)\qquad T = \frac{R}{[n + \alpha\,(l) + \alpha\,(l)/n^2]^2} = \frac{R}{n^{*2}},$$

where, for a given element, the constants $\alpha$ and $\alpha$ depend

[2] They are frequently also called *f′*-, *f″*-terms, *etc.*

*only* on the subordinate quantum number, so that their value is the *same* for all $s$-terms, the same for all $p$-terms, and so on. $R$ denotes, as usual, Rydberg's constant, and $n$ the principal quantum number. The first correction term $a$ is known as the *Rydberg correction*, and the second term $a$ as the *Ritz correction;* the number $n^*$ is called the *effective quantum number.*

As a first approximation the Rydberg correction is sufficient. It is always *negative*, as has been found from experiment, whereas the Ritz correction is always *positive.* Somewhat complicated considerations, which must be omitted here,[3] afford a theoretical basis for this fact, and they also yield formulæ for the correction terms which agree fairly well with the values experimentally obtained. These formulæ explain why, with increasing $l$, the correction terms decrease very rapidly.[4] In point of fact, the correction terms are quite insignificant for the $f$-terms, so that these exhibit a *complete analogy to hydrogen terms;* the deviations are small for the $d$-terms, but considerably larger for the $p$-terms, and generally very large for the $s$-terms. *Schrödinger* [5] has sought an explanation for this behaviour in the assumption that such terms correspond to so-called "*penetrating orbits*," a part of the orbit of the electron lying *within the atomic core*—the body of the atom less the valency electron. This assumption also explains why the correction terms increase in value with the atomic number.

Since the effective quantum numbers are experimentally known, and the theoretical corrections can be calculated at least approximately, the *true quantum numbers*, or values of $n$, can be derived for the corresponding terms. Thus the lowest $s$- and the lowest $p$-terms for lithium, which commences the second period in the system of the elements, are found to be two-quantum terms; for sodium they are three-quantum, for potassium four-, for rubidium five-, and for cæsium six-quantum terms. On the other hand the lowest $d$-term for *all* the alkali metals proves to be a three-quantum, and the lowest $f$-term a four-quantum term. Table VI reproduces the effective and, in brackets, the true quantum numbers for the lowest terms of the alkali

[3] *Cf.* Born's "*Vorlesungen über Atommechanik*" (Berlin: Springer, 1925), I. § 28.

[4] According to these formulæ, in which the question of the polarization of the atomic core plays an essential part, the Rydberg correction is inversely proportional to the fifth power, and the Ritz correction to the cube, of the subordinate quantum number.

[5] E. Schrödinger, *Zeitschr. f. Phys.*, **4** (1921), p. 347.

metals. It is evident from this table that, for all the alkali metals, the lowest term of all, the *fundamental term*, is an *s*-term. This is confirmed by the experimental fact that *only the principal series* occurs in the *absorption spectra* produced with *cold vapours* of these metals or their compounds (in which case the atoms have not been previously excited), and the principal series is the only one in which an *s*-term represents the final level on emission.

TABLE VI

*Effective and True Quantum Numbers of the Lowest Spectral Terms for the Alkali Metals*

| Alkali metal. | Lowest *s*-term. | Lowest *p*-term. | Lowest *d*-term. | Lowest *f*-term. |
|---|---|---|---|---|
| Lithium . | 1·59 (2) | 1·96 (2) | 3·00 (3) | 4·00 (4) |
| Sodium . | 1·63 (3) | 2·12 (3) | 2·99 (3) | 4·00 (4) |
| Potassium . | 1·77 (4) | 2·23 (4) | 2·85 (3) | 3·99 (4) |
| Rubidium . | 1·80 (5) | 2·28 (5) | 2·77 (3) | 3·99 (4) |
| Cæsium . | 1·87 (6) | 2·33 (6) | 2·55 (3) | 3·98 (4) |

It follows from the discussion in § 101 that, apart from the *s*-terms, all the terms of the alkali metals must be *doublets*. Experiment shows—and the result has a theoretical basis—that the *doublet separation* increases rapidly with the atomic number, but that it diminishes with increasing values of the subordinate and the effective quantum numbers.[6] Since the *principal series* arises from the combination of the lowest *s*-term with the several double *p*-terms, with which the separation diminishes as the principal quantum number increases, the *lines of the principal series* must represent *ever narrowing doublets*. The

[6] *Landé* has proposed the following formula for the doublet separation in the alkali metals :—

$$\Delta\nu = R\,\alpha^2\,\frac{Z_i^2\,Z_a^2}{n^{*3}\,l\,(l+1)}.$$

In this expression $\alpha$ denotes the constant introduced in § 104 eqn. (20) for the explanation of fine structure in the hydrogen spectrum ; $Z_a$ is to be given a value of unity for neutral atoms, and of two for simply ionized atoms ; and $Z_i$ denotes the mean effective nuclear charge for the inner portion of the orbit. *Cf.* E. Back and A. Landé, " *Zeeman-Effekt und Multiplett-Struktur* " (Berlin ; Springer, 1925).

series limit, determined by the simple $s$-term, is the same for both components of the doublets. On the other hand, the *sharp subordinate series* is due to combinations between $s$-terms and the lowest $p$-term, which itself represents a doublet, so that the lines of this series must be doublets with a *constant frequency difference*, and the series has a *double series limit.*

The *diffuse subordinate series* arises from combinations between double $p$-terms and double $d$-terms, *i.e.*, terms with the subordinate quantum number one and two respectively. The corresponding values of $j$ are hence $\frac{1}{2}$ and $\frac{3}{2}$, and $\frac{3}{2}$ and $\frac{5}{2}$. Since the inner quantum number $j$ can change only by 1 or 0, of the four conceivable combinations those between the $j$-values $\frac{1}{2}$ and $\frac{5}{2}$ are excluded. The lines of the diffuse subordinate series are therefore *triplets*, as also are those of the fundamental series, which is likewise due to combinations between pairs of doublet terms.

It is a peculiarity of the two components of the lowest $p$-level that from them, in the case of light emission, *only a return to the normal state is possible.* An atom which *absorbs* light of one of these two wave-lengths, and is thereby excited, can thus *emit* light only of the *same* wave-length on its return to the normal state. The double line corresponding to the combination of the lowest $p$-term with the fundamental term is therefore also known as the *resonance line* of the alkali metals. In the case of sodium it is the well-known yellow, so-called $D$-line ; this is the first line of the principal series, in which it has been possible to measure exactly no fewer than fifty-seven lines. In any spectrum there is, of course, not merely a single principal series, a single diffuse and a single sharp subordinate series, but there is a multiplicity of such series ; for example, the final level of a principal emission series may be any arbitrary $s$-term, and not only the fundamental term. By principal and subordinate series in the narrower senses of the words, however, we always understand the series of which the final levels on emission are the lowest $s$-level and the lowest $p$-doublet level respectively.

The first line in the principal series of the alkali metals obeys with great accuracy the relation derived in a previous section—§ 93 eqn. (3)—

$$\lambda V = 1{\cdot}234 \times 10^4, \tag{4}$$

where $V$ is the *excitation potential* measured in volts, and $\lambda$ is measured in Å. This equation is also satisfied when

$\lambda$ is given the reciprocal value of the fundamental term and $V$ the value of the *ionization potential*. As Table VII

TABLE VII

*Excitation and Ionization Potentials of the Alkali Metals (in Volts)*

| | | Li | Na | K | Rb | Cs |
|---|---|---|---|---|---|---|
| Excitation potential . | calc. | 1·84 | 2·09 | 1·61 | 1·6 | 1·4 |
| | obs. | — | 2·12 | 1·55 | 1·6 | 1·48 |
| Ionization potential . | calc. | 5·37 | 5·12 | 4·32 | 4·16 | 3·88 |
| | obs. | — | 5·13 | 4·1 | 4·1 | 3·9 |

shows, the directly *observed* values for the excitation and ionization potentials of the alkali metals agree well with those *calculated* from spectroscopic data.

The so-called *spark spectra of the alkaline earths, viz.*, the spectra emitted by the *ionized* atoms of these elements, exhibit a great similarity to the neutral "*arc spectra*" of the alkali metals.[7] In the spark spectra the arc terms appear *quadrupled* in magnitude, because an atomic core with the charge $(+2e)$ is here at work, as against one with a charge $(+e)$ in the case of the arc spectra of the alkali metals. A relation exists similar to that between the spectrum of ionized helium, previously described, and the hydrogen spectrum. Hence, if the spark terms of the alkaline earths be represented in the form $4R/n^{*2}$, a fair agreement is found between the effective quantum numbers of the alkaline earth ions thus obtained and those of the neutral alkali atoms.

Other analogues to the spectra of the alkalis are furnished by the *higher spark spectra* produced by atoms which are *several times ionized*, all the valency electrons except one being removed. A spectrum of this kind was first discovered by *Paschen*[8] in 1923, with doubly ionized *aluminium*,

[7] The spectra of ionized atoms presuppose higher temperatures or stronger electrical fields than the spectra of neutral atoms. Whereas the latter are generated in discharge tubes or, in particular, in the electric arc, spectra of ionized atoms are best produced in an electric spark discharge. Spectra of neutral and of ionized atoms are therefore commonly distinguished, according to the method of generation, as arc and spark spectra.

[8] F. Paschen, *Ann. der Phys.*, **71** (1923), p. 142.

and somewhat later *Fowler* [9] investigated the spectrum of trebly ionized *silicon*. As compared with the spectra of the alkalis, these exhibit nine-fold and sixteen-fold terms respectively. Finally, using the explosive vacuum spark method, *Millikan* and *Bowen*,[10] between 1924 and 1926, succeeded in obtaining the spectra of so-called "stripped atoms" (atoms with only one valency electron left) of all the elements in the second period up to *oxygen*, and in the third period up to *chlorine*.[11] Here also analogies with the spectra of the alkali metals were found. Furthermore, the systems of terms in the arc spectra of the elements copper, silver, and gold, are very similar to those in the spectra of the alkalies, all these elements having much in common in their chemical behaviour, and all belonging to the first group in the periodic system.

### § 107. The Spectra of Atoms with Two Valency Electrons.

When an atom contains several valency electrons, we have to deal in general with an equal number of *subordinate quantum numbers* $l_i$, and also with an additional subordinate quantum number $L$, which corresponds to the *resultant* of the individual vectors $\mathbf{l}_i$. The several values of $L$ ($L = 0, 1, 2, 3$, *etc.*) are associated with terms which may be denoted by capital letters, $S$-, $P$-, $D$-, $F$-terms, and so on. For example, a $P$-term ($L = 1$) may be determined by two $d$-electrons ($l_1 = 2, l_2 = 2$). Furthermore, a quantum number ($r_i = \frac{1}{2}$), corresponding to the *spin*, is to be ascribed to each valency electron. Composition of all the similarly or oppositely directed vectors $\mathbf{r}_i$ yields a resultant with an associated quantum number $r$, the value of which, for two valency electrons, must be zero or unity.

Finally, composition of the two vectors with the quantum numbers $L$ and $r$ gives the total angular momentum of the atom, and hence the *inner quantum number* $j$, the value of which determines the multiplicity of the terms. If we consider atoms with *two valency electrons*, we obtain thus a *singlet system* for the case $r = 0$, but a *triplet system* for the case $r = 1$, corresponding to the three possible $j$-values of $(L + 1)$, $L$, and $(L - 1)$. In the singlet system all the terms should be simple, whereas, in the triplet system, all of them would be triple, with the sole exception of the

[9] A. Fowler, *Proc. Roy. Soc. A.*, **103** (1923), p. 413.

[10] R. A. Millikan and I. S. Bowen, *Phys. Rev.*, 1924–6, *passim*.

[11] An example of such a stripped atom would be five-fold ionized oxygen, denoted by $O_{VI}$ (the neutral atom being $O_I$), or six-fold ionized chlorine ($Cl_{VII}$).

*S*-terms. In point of fact, a scheme of terms of this kind is formed by the elements in the second group of the periodic system, the *alkaline earths* together with *cadmium*, *zinc*, and *mercury*.

The spectral terms of these elements, with which we are longest acquainted, correspond throughout to the especially simple case in which one of the two valency electrons is an *s*-electron, so that $l_1$ may be equated to zero, the quantum numbers $L$ and $l_2$ thus coinciding. The scheme of terms obtained in this way may be regarded, in a narrower sense, as the system of terms for the element in question ; and, as we have just said, it resolves into a singlet and a triplet system. We may distinguish these by prefixing the index 1 or 3 to the capital letter denoting a term, *e.g.*, $^1S$, $^1P$, $^1D$ . . . $^3S$, $^3P$, $^3D$, *etc.*[1] The inner quantum number is written as a subscript, thus, $^3P_2$. In this way we obtain the following scheme of terms :—

$$\begin{array}{cccc} {}^1S_0 & {}^1P_1 & {}^1D_2 & {}^1F_3 \ . \ . \ . \\ & & {}^3S_1 & \\ {}^3P_0 & {}^3P_1 & {}^3P_2 & \\ {}^3D_1 & {}^3D_2 & {}^3D_3 & \\ {}^3F_2 & {}^3F_3 & {}^3F_4 \ \textit{etc.} & \end{array}$$

Where necessary, the principal quantum number can be inserted before the symbol, but generally it may be omitted. With simple terms, and when not otherwise required, the $j$-subscript may obviously be omitted.

The arrangement of the terms follows from the two rules, previously stated, that only those combinations are possible for which

(1) $$\Delta L = \pm 1, \quad \Delta j = \pm 1 \text{ or } 0.$$

It is found by experiment that, in addition to lines which would transgress either of these rules, lines corresponding to the transition

(2) $$(j = 0) \longrightarrow (j = 0)$$

also do not occur ; thus the line $^3P_0 \rightarrow {}^1S_0$ is missing. Combinations between the terms of the two systems, or so-called *intercombinations* (*i.e.*, combinations between singlets and triplets), produce *weaker* lines than those between

[1] The symbol $^3S$ is used for the simple *S*-term belonging to the triplet system ($r = 1$), because only in this way can a distinction be made between the *S*-terms of the two systems.

two singlet or two triplet terms. As an example of a term scheme, Fig. 62 reproduces the most important terms and lines in the *arc spectrum of mercury*. It is clearly shown that the $^1S$-term represents the *fundamental term*, and that there are two terms from which, on emission, only a return to the normal state is possible : these are the terms $^1P$ and $^3P_1$. To their combinations with the fundamental term correspond the two *resonance lines* of mercury which, with wave-lengths of 1,849 and 2,537 Å., lie far in the ultra-violet. We can calculate from them the two excitation potentials of mercury, and the values so derived are 4·87 and 6·67

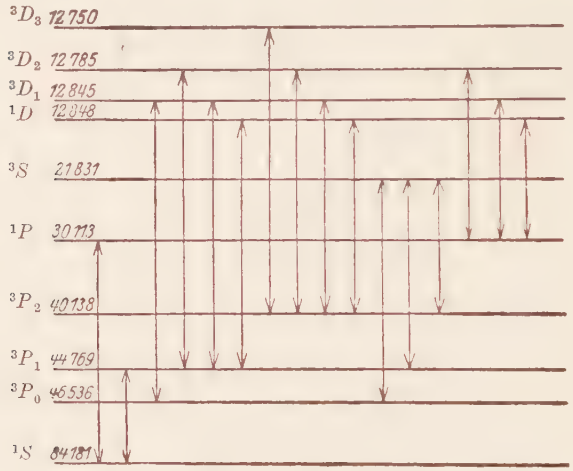



FIG. 62.—Scheme of the arc spectrum of mercury.

volts ; whereas the potentials actually observed by *Franck* and *Einsporn* [2] were 4·9 and 6·7 volts. A similar good agreement exists also for the excitation and ionization potentials of the other elements in the second period. Again, it is evident from the scheme of Fig. 62 that, in the multiplets of mercury, the component with the smallest value of $j$ represents the lowest term, and that successively larger values of $j$ represent successively higher terms. Multiplets of this kind, which are characteristic of all the elements in the first three periods, are called *regular multiplets*.

The spectral terms of the elements in the second period can also be represented by the *Rydberg-Ritz formula*, although

[2] *Cf.* § 93, footnote 1.

the corrections in general are larger than for the alkali metals. In the case of the $S$-, $P$-, and $F$- terms, the true quantum numbers are the same as for the corresponding alkali metals ; but, whereas for the latter the lowest $d$-term is always a three-quantum one, the lowest $D$-term for strontium proves to be a four-quantum, and that for barium a five-quantum term.

Besides the terms of the form $R/n^{*2}$, however, other terms also occur in the spectra of the alkaline earths, as was found in 1923 by *Wentzel* [3] ; these terms have the form

$$T = \frac{R}{n^{*2}} - (T_0' - T'), \tag{3}$$

where $T_0'$ denotes the *fundamental term in the spark spectrum* of the element concerned, and $T'$ a certain term in this spectrum. This remarkable relation was correctly ascribed by Wentzel to the correspondence of the terms represented by it to atomic states in which *two electrons* are *simultaneously excited*. A spectral line due to a combination of one of these terms with an ordinary term would thus be generated by the simultaneous transition of two electrons.

The terms which can be represented by eqn. (3) and, in some circumstances, may be *negative* [4] are called the *displaced terms* [5] *of the alkaline earths ;* they were detected first for *calcium*, and subsequently for the other elements in the second period. An interpretation of them in the sense of the vectorial model of the atom is due principally to *Russell* and *Saunders*, who succeeded in deriving the values of $l_i$ for the several displaced terms from the observed combinations and the selection rules [eqn. (1)]. If we represent a displaced term in the form $l_1\, l_2\, L$, using instead of numbers the corresponding symbols in the series $s, p, d, f \ldots$ $S, P, D, F$, *etc.*, and if we prefix the numeral expressing the multiplicity as an index to the capital letter, we may express the displaced terms hitherto discovered in the form [6]

$$d\,d\,{}^3P\,;\; d\,p\,{}^3P\,;\; d\,p\,{}^3D\,;\; d\,p\,{}^3F\,;\; d\,d\,{}^3D\,;\; p\,p\,{}^3P. \tag{4}$$

In the system of ordinary terms, one of the two electrons is

[3] The new terms were discovered about 1923 by *R. Goetze*, *A. del Campo*, and *Russell* and *Saunders*. Their existence was explained by Wentzel in the *Physikal. Zeitschr.*, 24 (1923), p. 104, and 25 (1924), p. 182.

[4] In such circumstances the energy required for the transition of an atom from the fundamental state into a doubly excited state is larger than the ionization energy.

[5] Also known as " primed " or " confused " terms.

[6] We may, of course, write either $dp^3P$ or $pd^3P$ : both symbols mean the same.

always an *s*-electron, so that the fundamental term may be written as follows :—

(5) Fundamental term (second period) $= s\ s\ ^1S$.

For the rest, in the case of the alkaline earths, *only intercombinations* of the displaced with ordinary terms are known : combinations between displaced terms have not been observed. Displaced terms have been discovered also in the *spark spectra* related to the spectra of the alkaline earths, *i.e.*, in those of the simply ionized atoms of the third group in the periodic system, or in the spectra produced by atoms, several times ionized, of the fourth, fifth, sixth, and seventh groups.[7]

Just as the hydrogen spectrum can be interpreted as a degenerate alkali spectrum, so also can the *arc spectrum of helium* be regarded as a *degenerate alkaline earth spectrum*. Helium exhibits two different systems of terms, and for a long time transitions between the two were thought to be impossible ; in 1922, however, *Lyman* succeeded in identifying such an *intercombination*.[8] Previously the two systems of terms had been ascribed to two different modifications of helium, one system being called the *parhelium* and the other the *orthohelium* system. The former consists of *singlet terms*, and contains a normal state in which *both electrons* are *one-quantum s-electrons*. The orthohelium system consists of *doublets* instead of triplets, as would be expected, but this phenomenon may possibly be due to a *coincidence of two triplet components*.[9] The fundamental term of orthohelium is also an *s*-term, which, however, must be assumed to be a *two-quantum* one.[10]

The remarkable antithesis between parhelium and orthohelium has been explained by *Heisenberg*[11] on the basis of *quantum mechanics*. The identity of the two electrons revolving round the nucleus produces a kind of *resonance*, as Heisenberg pointed out ; and, as he was also able to deduce, in consequence of this phenomenon of

[7] *Cf.* R. A. Millikan and I. S. Bowen, "*Series Spectra of Two-valence-electron Atoms of* $P_{IV}$, $S_V$, *and* $Cl_{VI}$," *Phys. Rev.*, **25** (1925), p. 591.

[8] T. Lyman, *Nature*, **110** (1922), p. 278 ; *Astrophys. Journ.*, **60** (1924), p. 1.

[9] *Cf.* W. Heisenberg, *Zeitschr. f. Phys.*, **39** (1926), p. 499.

[10] The earlier quantum theory completely failed in the calculation of the *ionization potential of helium*. On the basis of Schrödinger's theory, however, *G. W. Kellner* [*Zeitschr. f. Phys.*, **44** (1927), p. 91] succeeded in deriving a value which agreed well with experiment : a more accurate derivation by the same method was effected by *E. A. Hylleraas* [*Zeitschr. f. Phys.*, **48** (1928), p. 469]. Employing a different method, *J. C. Slater* [*Proc. Nat. Acad. Sc., U.S.A.* **13** (1927), p. 423], obtained a similar result.

[11] W. Heisenberg, *Zeitschr. f. Phys.*, **41** (1927), p. 239.

resonance, the *term system* corresponding to the *coupling* of the two electrons, is *resolved into partial systems*, between which only *weak intercombinations* occur as a result of the electronic spin. By a generalization of these considerations, Heisenberg was also able to explain why the *electron-magnets* should always take up *parallel* positions in the *same* or the *opposite* sense in an atom. He also furnished a basis for the "*cosine law*" mentioned in a previous section (§ 101), on which rests the *interval rule* for multiplets.

### § 108. The Spectral Terms in Equivalent Systems of Electrons.

If we wish to use the relations of the alkali and alkaline earth spectra, which we have previously discussed, as a basis of investigation of the optical spectra of the other elements, we must first answer two quite general questions. The first

TABLE VIII

*Scheme of Terms for Two Equivalent p-Electrons*

| $m_1\,(l)$ | $m_2\,(l)$ | $m_1\,(r)$ | $m_2\,(r)$ | $m_L$ | $m_r$ |
|---|---|---|---|---|---|
| 1 | 1 | $+\frac{1}{2}$ | $-\frac{1}{2}$ | 2 | 0 |
| 1 | 0 | $\pm\frac{1}{2}$ | $\pm\frac{1}{2}$ | 1 | 1 0 0 −1 |
| 1 | −1 | $\pm\frac{1}{2}$ | $\pm\frac{1}{2}$ | 0 | 1 0 0 −1 |
| 0 | 0 | $+\frac{1}{2}$ | $-\frac{1}{2}$ | 0 | 0 |
| 0 | −1 | $\pm\frac{1}{2}$ | $\pm\frac{1}{2}$ | −1 | 1 0 0 −1 |
| −1 | −1 | $+\frac{1}{2}$ | $-\frac{1}{2}$ | −2 | 0 |

question is: what terms are possible for a *system of equivalent electrons?* The second question concerns the changes effected in a term, peculiar to a given system of electrons, on the *addition of another given electron.*

An answer to the first question is provided by *Pauli's principle.* It follows from this principle, as was explained in a previous section (§ 103), that there cannot be more than *two* equivalent *s*-electrons, and that, since in a system of two equivalent *s*-electrons $L = r = 0$, only a $^1S$-term is possible.

With equivalent *p*-electrons, we have to distinguish six cases according as their number is 1, 2, 3, 4, 5 or 6. A single *p*-electron gives, of course, a *P*-term, and that a *doublet term*, corresponding to the two possibilities of opposite

arrangements of the rotating electron. With *two* $p$-electrons we may have (using the notation of § 103)

$$m_1(l) \text{ or } m_2(l) = 1, 0, -1;$$
$$m_1(r) \text{ or } m_2(r) = +\tfrac{1}{2}, -\tfrac{1}{2}.$$

It must be remembered, however, that, according to Pauli's principle, when $m_1(l) = m_2(l)$, $m_1(r)$ must *differ* from $m_2(r)$. Altogether, we find for the pairs $m_L'[=\sum m_i(l)]$ and $m_r[=\sum m_i(r)]$ the possibilities contained in Table VIII.[1]

Since, for a given value of $L$, $m_L$ may have all values from $(+L)$ to $(-L)$, and for a given value of $r$, $m_r$ may have all values from $(+r)$ to $(-r)$, we see that Table VIII contains the following three $L$ - $r$ - combinations :—

(1) $\quad L = 2, r = 0;\ L = 1, r = 1;\ L = 0, r = 0.$

It follows, therefore, from the discussion in § 101 on the multiplicity of terms that a system of two equivalent $p$-electrons yields the terms $^1D, {}^3P, {}^1S$. Now it is found by experience—and Heisenberg has afforded a theoretical proof [2]—that the position of the terms in a system of equivalent electrons is governed by a rule which admits of no known exception. This rule states that, the *greater the multiplicity of a term*, the *lower* it lies, while, among terms of equal multiplicity, the one with the *larger value* of $L$ lies the *lower*. Hence the three terms are to be arranged in order of falling levels : $^1S, {}^1D, {}^3P$.

## Table IX

*Scheme of Terms for Three Equivalent p-Electrons*

| $m_1(l)$ | $m_2(l)$ | $m_3(l)$ | $m_1(r)$ | $m_2(r)$ | $m_3(r)$ | $m_L$ | $m_r$ |
|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | $+\frac{1}{2}$ | $-\frac{1}{2}$ | $\pm\frac{1}{2}$ | 2 | $\pm\frac{1}{2}$ |
| 1 | 1 | −1 | $+\frac{1}{2}$ | $-\frac{1}{2}$ | $\pm\frac{1}{2}$ | 1 | $\pm\frac{1}{2}$ |
| 1 | 0 | 0 | $\pm\frac{1}{2}$ | $+\frac{1}{2}$ | $-\frac{1}{2}$ | 1 | $\pm\frac{1}{2}$ |
| 1 | 0 | −1 | $\pm\frac{1}{2}$ | $\pm\frac{1}{2}$ | $\pm\frac{1}{2}$ | 0 | $\pm\frac{1}{2}\pm\frac{1}{2}\pm\frac{1}{2}\pm\frac{3}{2}$ |
| 1 | −1 | −1 | $\pm\frac{1}{2}$ | $+\frac{1}{2}$ | $-\frac{1}{2}$ | −1 | $\pm\frac{1}{2}$ |
| 0 | 0 | −1 | $+\frac{1}{2}$ | $-\frac{1}{2}$ | $\pm\frac{1}{2}$ | −1 | $\pm\frac{1}{2}$ |
| 0 | −1 | −1 | $\pm\frac{1}{2}$ | $+\frac{1}{2}$ | $-\frac{1}{2}$ | −2 | $\pm\frac{1}{2}$ |

[1] The values given for the first and second electron can, of course, be interchanged without any new combination arising : such duplicate combinations, due only to an interchange, are therefore omitted in Tables VIII to XI.

[2] *Vide* § 107, footnote 10.

For a system of *three equivalent* $p$-electrons we obtain, on the basis of Pauli's principle, the possibilities reproduced in Table IX. From this table we see that there are again three possible combinations of $L$ and $r$, *viz.*,

(2) $\quad L = 2, r = \frac{1}{2}\,;\ L = 1, r = \frac{1}{2}\,;\ L = 0, r = \frac{3}{2}\,.$

The corresponding terms are $^2D, {}^2P, {}^4S$, or, arranged in order of falling levels, $^2P, {}^2D, {}^4S$.

The combinations obtained in a similar way for a system of *four equivalent* $p$-electrons, on the basis of Pauli's principle, are contained in Table X. The $L$ - $r$ - combinations

TABLE X

*Scheme of Terms for Four Equivalent p-Electrons*

| $m_1\,(l)$ | $m_2\,(l)$ | $m_3\,(l)$ | $m_4\,(l)$ | $m_1\,(r)$ | $m_2\,(r)$ | $m_3\,(r)$ | $m_4\,(r)$ | $m_L$ | $m_r$ |
|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | $+\frac{1}{2}$ | $-\frac{1}{2}$ | $+\frac{1}{2}$ | $-\frac{1}{2}$ | 2 | 0 |
| 1 | 1 | 0 | −1 | $+\frac{1}{2}$ | $-\frac{1}{2}$ | $\pm\frac{1}{2}$ | $\pm\frac{1}{2}$ | 1 | 1 0 0 −1 |
| 1 | 1 | −1 | −1 | $+\frac{1}{2}$ | $-\frac{1}{2}$ | $+\frac{1}{2}$ | $-\frac{1}{2}$ | 0 | 0 |
| 1 | 0 | 0 | −1 | $\pm\frac{1}{2}$ | $+\frac{1}{2}$ | $-\frac{1}{2}$ | $\pm\frac{1}{2}$ | 0 | 1 0 0 −1 |
| 1 | 0 | −1 | −1 | $\pm\frac{1}{2}$ | $\pm\frac{1}{2}$ | $+\frac{1}{2}$ | $-\frac{1}{2}$ | −1 | 1 0 0 −1 |
| 0 | 0 | −1 | −1 | $+\frac{1}{2}$ | $-\frac{1}{2}$ | $+\frac{1}{2}$ | $-\frac{1}{2}$ | −2 | 0 |

are found to be the same as for two equivalent electrons, *viz.*,

(3) $\quad L = 2, r = 0\,;\ L = 1, r = 1\,;\ L = 0, r = 0\,.$

Hence we obtain again the same three terms $^1D, {}^3P, {}^1S$ ; or, in order, $^1S, {}^1D, {}^3P$.

TABLE XI

*Scheme of Terms for Five Equivalent p-Electrons*

| $m_1\,(l)$ | $m_2\,(l)$ | $m_3\,(l)$ | $m_4\,(l)$ | $m_5\,(l)$ | $m_1\,(r)$ | $m_2\,(r)$ | $m_3\,(r)$ | $m_4\,(r)$ | $m_5\,(r)$ | $m_L$ | $m_r$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 0 | 0 | −1 | $+\frac{1}{2}$ | $-\frac{1}{2}$ | $+\frac{1}{2}$ | $-\frac{1}{2}$ | $\pm\frac{1}{2}$ | 1 | $\pm\frac{1}{2}$ |
| 1 | 1 | 0 | −1 | −1 | $+\frac{1}{2}$ | $-\frac{1}{2}$ | $\pm\frac{1}{2}$ | $+\frac{1}{2}$ | $-\frac{1}{2}$ | 0 | $\pm\frac{1}{2}$ |
| 1 | 0 | 0 | −1 | −1 | $\pm\frac{1}{2}$ | $+\frac{1}{2}$ | $-\frac{1}{2}$ | $+\frac{1}{2}$ | $-\frac{1}{2}$ | −1 | $\pm\frac{1}{2}$ |

With *five equivalent* $p$-electrons the possibilities are much

restricted by Pauli's principle, as Table XI shows, and we find only one possible combination :—

$$L = 1, r = \tfrac{1}{2}, \tag{4}$$

or a $^2P$-term, as in the case of a single electron. It follows directly from the considerations of § 103 that, for a system of *six* equivalent $p$-electrons,

$$L = r = 0 ; \tag{5}$$

thus we obtain a $^1S$-term.

The results of our considerations of systems with equivalent $p$-electrons are collected in Table XII. The terms

TABLE XII

*Terms in a System of Equivalent p-Electrons*

| Number of equivalent electrons. | Spectral terms. |
|---|---|
| 6 | $^1S$ |
| 1 or 5 | $^2P$ |
| 2 or 4 | $^1S$ $^1D$ $^3P$ |
| 3 | $^2P$ $^2D$ $^4S$ |

in a system of equivalent $d$-electrons can be calculated in like manner ; the results are naturally considerably more complicated,[3] and obviously those for systems of equivalent $f$-electrons will be even more so.

Finally, we have still to answer the question in what way does a given spectral term, produced by a system of equivalent electrons, change when a *further, non-equivalent, electron* is *added* to the system. Let $l$ be the subordinate quantum number of the equivalent electrons and $l'$ that of the additional electron ($l$ and $l'$ may be identical if the principal quantum number differs) ; let $L$ be the subordinate quantum number of the term due to the equivalent electrons and $x$ its multiplicity. New terms will then be associated with all the values for $L$ which lie between the sum and the magnitude of the difference of $L$ and $l'$. Thus we have

$$L' = (L + l'), (L + l' - 1) \ldots |L - l'| . \tag{6}$$

[3] *Cf.* F. Hund's "*Linienspektren und periodisches System*" (Berlin ; Springer, 1927).

The multiplicity must be increased or decreased by unity, so that

$$x' = x \pm 1. \tag{7}$$

### § 109. Spectra of Elements in the Short Periods.

Now that we have dealt with the spectra of hydrogen, of helium, and of the elements in the first two groups of the periodic system, we are in a position to interpret the term structure of optical spectra in general, with the help of the rules which we have just derived. It follows necessarily from *Pauli's principle* that the *first period* of the system of the elements comprises *only two elements*, since no atom can contain more than two one-quantum electrons. Since, according to the same principle, the maximum number of *two-quantum* electrons is *eight*, and since the second period actually includes eight elements, we may reasonably assume that, in the normal state, all the eight elements in the second period contain only two-quantum electrons, apart from two one-quantum electrons. In the case of neon, which concludes the second period, two of the eight two-quantum electrons obviously make up a complete *s*-sub-group, and six a complete *p*-sub-group. Now we know that the fundamental term of *lithium* is a two-quantum *s*-term, and we know likewise that *sodium*, which commences the third period also of eight elements, has a three-quantum *s*-term as its fundamental term. The lowest term of the next element after sodium, *magnesium*, is again a three-quantum *s*-term, whereas that of the following element, *aluminium*, is found to be a three-quantum *p*-term. These spectroscopic results show that the following arrangement of the two- and three-quantum electrons in the *eight groups* of the second and third periods is probable :—

$$(1) \quad \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ s & s^2 & s^2\,p & s^2\,p^2 & s^2\,p^3 & s^2\,p^4 & s^2\,p^5 & s^2\,p^6. \end{matrix}$$

In this formula the symbol $s^2\,p^6$ denotes a group of two *s*- and six *p*-electrons, *etc.*

In point of fact, the conclusions to be derived from this scheme agree well with experiment. We have already discussed the alkali metals and the alkaline earths in previous sections. Since $s^2$ represents a completed sub-group which has no influence on the vectorial frame of the atom, we may expect a system of doublet terms for *aluminium*, with a three-quantum $^2P$-term as fundamental term. Actually, aluminium exhibits a very simple *system of doublet*

*terms*, which resembles that of the alkali metals to a considerable extent.

According to the above scheme, the normal arrangement for the *arc spectrum of silicon* [1] would be $s\,s\,p\,p$; we see from Table XII that the corresponding spectral terms are $^1S$, $^1D$, $^3P$. Besides the normal state, we might also have an arrangement in which the fourth electron is an $s$-electron: its principal quantum-number must be at least *four*, because, by Pauli's principle, there is no room in the silicon atom for any more one-, two-, or three-quantum $s$-electrons. Since a single $p$-electron yields a $^2P$-term, it follows from § 108 eqns. (6) and (7), that the addition of an $s$-electron to the $p$-electron will produce the terms $^1P$ and $^3P$. Another possible arrangement would be $s\,s\,p\,d$. The addition of a $d$-electron ($l = 2$) to a system associated with a $^2P$-term yields, according to § 108 eqns. (6) and (7), the following terms (in order of falling levels): $^1P$, $^1D$, $^1F$, $^3P$, $^3D$, $^3F$. The term scheme of neutral silicon, as *experimentally derived*, is reproduced in Fig. 63, in which is also shown the interpretation of the terms on the basis of the theoretical results just obtained.[2] The features of the arc spectrum of silicon are very similar to those both of the arc spectra of tin and lead, and of the first spark spectrum of nitrogen.

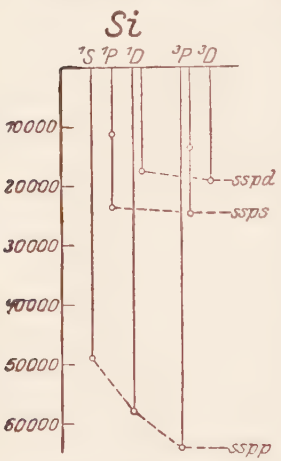



Fig. 63.—Term scheme of silicon (Hund, *l.c.*)

The general scheme of terms for an element in the *fifth group* is to be derived from the terms corresponding to a system of two equivalent $p$-electrons. According to Table XII, the *normal state* of such an element ($s\,s\,p\,p\,p$) will be associated with the terms $^2P$, $^2D$, $^4S$. The *fundamental term* is therefore an $S$-term belonging to a quartet system.

Simpler conditions may be expected for the elements in the *sixth group*. According to § 108 eqns. (6) and (7), on addition of a further electron the lowest term for the

[1] This has been investigated principally by *McLennan* and by *Fowler*.

[2] A further possible arrangement would be that in which a *non-equivalent* $p$-electron is added to the $s\,s\,p$-group. In this case we find simple and triple $S$-, $P$-, and $D$-terms, some of which have been observed in the spark spectrum of nitrogen.

fifth group, the $^4S$-term, will yield *triplet* and *quintet terms*: for an *s*-electron these will be *S*-terms, for a *p*-electron *P*-terms, and for a *d*-electron *D*-terms. Actually, triplet and quintet systems of these kinds are observed for both *oxygen* and *sulphur*.

For the elements in the seventh group, the *halogens*, we must again expect more complex series relationships, and little is experimentally known of these. On the other hand, the arc spectra of the *rare gases neon* and *argon* (in the eighth group) have been investigated with fair accuracy.[3] Their very complicated spectra can be interpreted with the help of theoretical predictions, as can the *spark spectra* of the *alkali metals*,[4] which have the characteristics of a rare gas spectrum.

## § 110. Spectra of Elements in the Long Periods.

The spectra of the elements in the fourth and fifth periods, each of which comprises *eighteen elements*, are considerably more complicated than those of the second and third periods. The third period ends with argon, but the three-quantum group of electrons, which may contain eighteen, is not then complete. For all the elements of this period in the normal state, the three-quantum group contains *only s-* and *p*-electrons, *d*-electrons being entirely *lacking*.

The element succeeding argon, *potassium*, proves to be an alkali metal, with a four-quantum *s*-term as its fundamental term. The next elements are *calcium* and *scandium*. Calcium is an alkaline earth, and in both its arc and its spark spectrum the fundamental term is a four-quantum *S*-term, as might well be expected also for the spectrum of *doubly ionized scandium*. Actually, however, as *Gibbs* and *White* [1] succeeded in demonstrating, and as *Bohr* had previously conjectured,[2] the fundamental term of $Sc^{++}$ is a *three-quantum D-term*. Hence we must conclude that the single electron, which exists in the doubly ionized scandium atom outside the argon core, is a *d*-electron. In the case of $Sc^{+}$ an additional electron is present, and

[3] The spectrum of neon was first analyzed by *Paschen*, and that of argon by *Saunders* and *Meissner*.

[4] The spark spectra of the alkalis were discovered by *Goldstein*. An elucidation of the terms in the spectrum of $K^{+}$ was first afforded by *de Bruin*.

[1] R. C. Gibbs and H. E. White, *Proc. Nat. Acad. U.S.A.*, **12** (1926), p. 598.

[2] *Vide* § 112.

this might equally well be an $s$- or a $d$-electron. The corresponding lowest terms, as derived from § 108 eqns. (6) and (7), or from considerations analogous to those underlying Table VIII, would be $^3D$ and $^3F$ respectively. Since the former alternative is the one experimentally confirmed, the arrangement $s\,d$ must be ascribed to $Sc^+$.

With neutral scandium either of the arrangements $s\,s\,d$ and $s\,d\,d$ is conceivable. In the first case the *fundamental term* would be $^2D$; in the second case $^4F$. Actually the first possibility is realized in the arc spectrum of scandium, and the second in the spark spectrum of the next element, *titanium*. In like manner we can derive the lowest terms for the succeeding elements on the assumption that, outside the rare gas core, 0, 1, or 2 $s$-electrons are present. When dealing with the $x$th element in the period we have therefore to distinguish between the three possible arrangements:

$$d^x,\ s\,d^{x-1},\ s^2d^{x-2}.$$

To these there correspond in general three different fundamental terms, so that, if the fundamental term be *experimentally* known, we can conversely derive from it the actual *arrangement of the electrons*. The results thus obtained are contained in Table XIII.

TABLE XIII

*Fundamental Terms and Arrangement of Electrons in the Long Periods*

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|
| K | Ca | Sc | Ti | V | Cr | Mn | Fe | Co | Ni | Cu |
| $^2S$ | $^1S$ | $^2D$ | $^3F$ | $^4F$ | $^7S$ | $^6S$ | $^5D$ | $^4F$ | $^3F$ | $^2S$ |
| $s$ | $s^2$ | $s^2\,d$ | $s^2\,d^2$ | $s^2\,d^3$ | $s\,d^5$ | $s^2\,d^5$ | $s^2\,d^6$ | $s^2\,d^7$ | $s^2\,d^8$ | $s\,d^{10}$ |
| Rb | Sr | Y | Zr | Cb | Mo | Ma | Ru | Rh | Pd | Ag |
| $^2S$ | $^1S$ | $^2D$ | $^3F$ | $^6D$ | $^7S$ | — | $^5F$ | $^4F$ | $^1S$ | $^2S$ |
| $s$ | $s^2$ | $s^2\,d$ | $s^2\,d^2$ | $s\,d^4$ | $s\,d^5$ | — | $s\,d^7$ | $s\,d^8$ | $d^{10}$ | $s\,d^{10}$ |

As we can see from this table,[3] in the fourth period

[3] The derivation of the electronic arrangement from the fundamental term is not entirely certain only in the case of Co (for which $s\,d^8$ would also be

we meet with arrangements containing both one and two $s$-electrons. In the fifth period distributions with a single $s$-electron are more common than in the fourth period. In the so-called *palladium series* (yttrium to palladium), the $d$-electrons are therefore apparently *more firmly bound to the nucleus* than in the so-called *iron series* (scandium to nickel). The behaviour of *palladium* is very remarkable, for, although it belongs to the fifth period, it does *not* contain *a single five-quantum electron* in the normal state. The elements succeeding copper in the fourth and silver in the fifth period are, both in chemical and in spectroscopical respects, similar to the corresponding elements in the second to the eighth groups in the short periods.

On the basis of the electronic arrangements reproduced in Table XIII. it is, of course, also possible conversely to calculate the higher terms of a given element, and thus to arrive at a theoretical interpretation of the observed arc and spark spectra. The elucidation of the extremely complicated spectra of the long periods is principally due to *Catalán* (spectrum of manganese), *Walters* and *Laporte* (iron spectrum), *Sommerfeld*, *McLennan*, and *Hund* (to whose treatise [4] reference should be made for details).

The spectroscopic behaviour of *copper* is noteworthy, because it exhibits *two different spectra*. One of these, an *alkali metal spectrum in character*, with a $^2S$ fundamental term, corresponds to the electronic distribution $s\,d^{10}$, while the other,[5] with a $^2D$ fundamental term, corresponds to the arrangement $s^2\,d^9$. This phenomenon is probably connected with the fact that copper occurs both in a univalent and a *bivalent* state. With *silver*, the lines of the second spectrum are not known, and silver is always univalent, perhaps because the $d$-electrons are more firmly bound to the nucleus in the fifth period than in the fourth. On the other hand, both univalent and trivalent *gold* are known, and, in addition

possible) and of Rh (where it might also be $s^2\,d^7$). For the arguments in favour of the distributions given in Table XIII *cide* Hund, *l.c.* The known fundamental terms and arrangements of simply *ionized* atoms are as follows :—

| | | | | | |
|---|---|---|---|---|---|
| $Ca^+$ | $^2S$ | $s$ | $Fe^+$ | $^6D$ | $s\,d^6$ |
| $Sc^+$ | $^3D$ | $s\,d$ | $Sr^+$ | $^2S$ | $s$ |
| $Ti^+$ | $^4F$ | $s\,d^2$ | $Y^+$ | $^3D$ | $s\,d$ |
| $V^+$ | $^5D$ | $d^4$ | $Zr^+$ | $^4F$ | $s\,d^2$ |
| $Cr^+$ | $^6S$ | $d^5$ | $Cb^+$ | $^5D$ | $d^4$ |
| $Mn^+$ | $^7S$ | $s\,d^5$ | | | |

[4] Hund, *l.c.*

[5] *Cf.* L. A. Sommer, *Zeitschr. f. Phys.*, **39** (1926), p. 711.

to the alkali-like spectrum, terms of a second spectrum have been found.

TABLE XIV

*Multiplicity of Spectral Terms in the Third Period*

| K | Ca | Sc | Ti | V | Cr | Mn | Fe | Co | Ni | Cu | Zn |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 2 | 1, 3 | 2, 4 | 1, 3, 5 | 2, 4, 6 | 3, 5, 7 | 4, 6, 8 | 3, 5, 7 | 2, 4, 6 | 1, 3,5 | 2, 4 | 1, 3 |
| $Ca^+$ | $Sc^+$ | $Ti^+$ | $V^+$ | $Cr^+$ | $Mn^+$ | $Fe^+$ | | | | | |
| 2 | 3 | 2, 4 | 3, 5 | 4, 6 | 7 | 6 | | | | | |

The *multiplicities* observed in the arc and spark spectra of the elements from potassium to zinc are reproduced in Table XIV. We may see in it a confirmation of the rule that the maximum multiplicity can never be greater than one more than the number of valency electrons. The table also shows that, as was discovered by *Sommerfeld* and expressed in his "*alternation rule*," *even* and *odd multiplicity* alternate, the former being found with an odd, and the latter with an even, *number of electrons*. The multiplets prove to be *regular*, when a lower term corresponds to a smaller value of $j$, or *inverted*, when the lower term is associated with a larger $j$-value, or *partially inverted*, when the succession of terms does not correspond to that of the $j$-values. As regards the *fundamental terms*, a *rule* established by *Slater* [6] is of universal validity: this rule states that, in the development of an equivalent group, the fundamental terms for the first members of the group are regular, for the last members inverted, and for the central members $S$-terms. In point of fact, the fundamental term for Sc, Ti, and V is regular; that for Fe, Co, and Ni is inverted; whilst for Cr and Mn (*cf.* Table XIII) it is an $S$-term which, since this is always simple, may be regarded as either regular or inverted.

The spectral lines themselves obey in general a *combination rule* which was propounded by *Laporte* [7] and by *Russell*.[8] According to this rule, only such combinations occur as give an *odd* number for the *sum* of the changes in the values of the $l_i$. Furthermore, there are only a few exceptions to a special rule, formulated by *Heisenberg*,[9] which states

[6] J. C. Slater, *Phys. Rev.*, **28** (1926), p. 291.
[7] O. Laporte, *Zeitschr. f. Phys.*, **23** (1924), p. 135, and **26** (1924), p. 1.
[8] H. N. Russell, *Science*, **69** (1924), p. 512.
[9] W. Heisenberg, *Zeitschr. f. Phys.*, **32** (1925), p. 841.

that the only permissible transitions are those in which, for one electron, $l_1$ changes by $\pm 1$, whereas, for *another* electron, $l_2$ changes by 0 or $\pm 2$.

Finally, as regards the spectra and electronic distributions in the *sixth period*, we can form three series out of the thirty-two elements in this period. The first of these is a series of eight, composed of the first two and the last six elements, *i.e.*,

Cs, Ba, Tl, Pb, Bi, Po, (85), Rn.

These elements correspond in a spectroscopic respect fairly closely to those in the eight groups of the second or third period. Of the first and the last nine of the remaining twenty-four elements we form the so-called *platinum series*, analogous to the iron and the palladium series :

La, Hf, Ta, W, Re, Os, Ir, Pt, Au, Hg.

We know that tungsten has a $^5D$ fundamental term, and this points to an $s^2 d^4$ distribution. Platinum has a $^3F$ fundamental term, indicating an $s^2 d^8$ arrangement. With gold and mercury the fundamental term and distribution are the same as with silver and cadmium.

Between lanthanum and hafnium lie *fourteen* very similar elements, the so-called *rare earths*. *Bohr* has explained their peculiar position by assuming that, after lanthanum, the further development of the five-quantum *d*-group and the six-quantum *s*-group is interrupted by the completion, between lanthanum and hafnium, of the four-quantum *f*-group of fourteen electrons. Whether, in addition to this group, the rare earths contain two five-quantum *d*-electrons and one six-quantum *s*-electron, or, conversely, one *d*-electron and two *s*-electrons, is not yet definitely known ; possibly both these distributions occur. In either case, however, very complicated spectra are to be foreseen. In the first case there would be fundamental terms with multiplicities up to 11 and *L*-values up to 9 ; even in the second case we must expect nine-fold multiplicity and *L*-values up to 8. We can look for simpler conditions only in the *three-fold ions of the rare earths*, since these obviously contain, outside the rare gas core, only four-quantum *f*-electrons, from one to fourteen in number.

The reduction of a line-spectrum to its terms is generally a most difficult problem, which, however, has already apparently been solved for very many elements. The first

step towards such a solution consists always in the discovery of *constant frequency differences between groups of lines*, whereby the definition of terms is rendered possible. With the help of the *combination rules* and Landé's *interval rule*, we can then calculate the values of the *inner quantum number* for the individual term components. On this basis the spectrum can be arranged in *multiplets*. As we shall see in

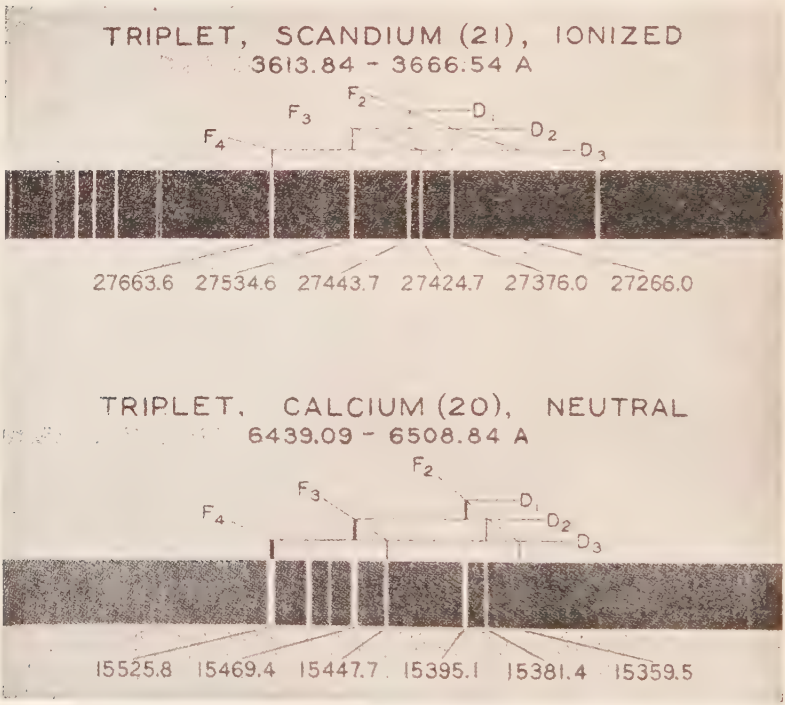



Fig. 64.—Line sextet (combination of two triplet terms) in the spectra of $Sc^+$ and Ca (after Meggers).

subsequent sections, the analysis of spectral terms is also facilitated by researches on the magnetic resolution of lines and on their intensities.

Fig. 64 is the reproduction of a photograph by *Meggers* [10] of a *line sextet*, on the one hand in the *spark spectrum of scandium*, and on the other hand in the *arc spectrum of calcium*, due to the combination of a *triplet D-term* with a

[10] *Cf. Journ. Opt. Soc. America*, 1924–5.

*triplet F-term.* A similar photograph, also by Meggers, is reproduced in Fig. 65: this shows a group of *fourteen lines*, in the *spark spectrum of chromium* and in the *arc spectrum of vanadium*, arising from a combination of a *quintet D-term* with a *sextet F-term.*




FIG. 65.—Line multiplet of fourteen components (combination of a quintet and a sextet term) in the spectra of Cr+ and V (after Meggers).

## § 111. X-ray Spectra.

In 1905 *Barkla* made the important discovery of *characteristic X-rays.* As Barkla observed, when primary X-rays impinge on a body, the *secondary* X-radiation emitted by it consists chiefly of *homogeneous rays*, the *hardness* of which is found to be a *characteristic property of the atoms*

of the body. Barkla measured the hardness of the rays by means of their *absorption* in aluminium plates, the absorption coefficient representing a constant characteristic of the element in question. He found the radiation from chemical compounds to be composed of the rays of the elements contained in the compounds. He further observed that the hardness of the characteristic rays *increases with the atomic weight*, and, moreover, that the majority of the elements appeared to possess not merely one, but *two* absorption coefficients differing widely in value. Barkla therefore distinguished between a harder *K-radiation* and a softer *L-radiation*. He attributed quite rightly to the limitations of experimental methods the fact that only one species of rays could be demonstrated in the case of many elements. For, on the one hand, the *K*-rays become so hard with increasing atomic weight that, in the state of X-ray technique at that time, it was impossible to excite them; on the other hand, the *L*-radiation becomes so soft with decreasing atomic weight that it could not be demonstrated by the methods then known. Furthermore, as Barkla discovered, characteristic rays are generated only by primary rays which are *harder* than the characteristic rays themselves.

The invention of the method of the *rotating crystal* (to be discussed subsequently) rendered possible an exact analysis of the characteristic radiation, and led to its resolution into a *line spectrum*, the exact measurement of *wave-lengths* taking the place of the inaccurate determination of hardness. We owe the foundation of *X-ray spectroscopy* to *Moseley* who, in 1913, by means of a specially constructed vacuum spectrograph, resolved the *K*- and the *L*-radiations of numerous elements into distinct sharp lines, the wave-lengths of which he determined accurately. Moseley found that both the *K*- and the *L*-series are of the same type for the different elements. In each series one line was found to be particularly strong, and this line he called the $K_\alpha$- and the $L_\alpha$-line respectively. More recent X-ray spectroscopic researches have resulted both in the observation of new lines in both series, and also in the discovery of two further series. These, on account of their softness, have actually been demonstrated only in elements of high and the highest atomic weights, and are known as the *M*- and the *N*-series. The *M*-series was discovered by *Siegbahn* and investigated principally by him and *Stenström*. The *N*-series was first detected by *Dolejsek*.

In laying the foundations of X-ray spectroscopy,

*Moseley* discovered an important relation which permits an *interpretation of the lines in X-ray spectra* from the point of view of the *quantum theory*. This relation expresses that the *frequency* of the strongest line in the $K$-series, the $K_\alpha$-*line*, can be represented by the formula

$$\frac{\nu}{c}(K_\alpha) = \frac{3}{4} R (z - a)^2, \tag{1}$$

where $R$ denotes the Rydberg constant, $z$ the nuclear charge number of the element concerned, and $a$ is a constant with a value differing but little from 1·6 for the majority of the elements.

Eqn. (1) finds an obvious explanation if we consider the $K_\alpha$-line as being due to a *transition of an electron from a two-quantum to a one-quantum state* [inasmuch as 3/4 is equal to $(1/1^2 - 1/2^2)$]. Provided, however, that we regard the atom, in rough approximation, as a hydrogen-analogue, the number $a$ is to be interpreted as a so-called *screening-constant*. The nucleus behaves as though a part of its charge, *viz.*, $a\,e$, were screened off; for not only is the electron that undergoes the transition attracted by the atomic nucleus, but it is also *repelled* by the other electrons which, as well as it, surround the nucleus. Since, however, the screening constant of the $K_\alpha$-line has almost the same value for the most various elements, this obviously means that the transition giving rise to the $K_\alpha$-line must occur in a region of the atom which is *similarly constituted* for all atoms of arbitrary nuclear charge numbers (provided these are not too small).

For the strongest line in the $L$-series, the so-called $L_\alpha$-line, Moseley likewise found the relation

$$\frac{\nu}{c}(L_\alpha) = \frac{5}{36} R (z - b)^2, \tag{2}$$

where, again, $b$ is a screening constant. Its value for most elements is about 3·5, but the variations in the values of the screening constant for the $L_\alpha$-line are considerably greater than in the case of the $K_\alpha$-line. Eqn. (2) may evidently be interpreted as meaning that the $L_\alpha$-line arises from a transition of an electron *from a three-quantum to a two-quantum state*. The second strongest line in the $K$-series, generally known as the $K_\beta$-line, is due to a transition from a three-quantum to a one-quantum state, as follows from its wave-length. In like manner, the second strongest line in the $L$-series, denoted by $L_\beta$,[1] arises from a transition

[1] Formerly this line was also denoted by $L_\gamma$.

from a four-quantum to a two-quantum state, and the strongest line in the $M$-series, the so-called $M_\alpha$-line, from a transition from a four-quantum to a three-quantum state. The transition from the four-quantum to the one-quantum state gives rise to a weak line in the $K$-series, the so-called $K_\gamma$-line, the wave-length of which is shorter than that of the $K_\beta$-line. A schematic representation of the origin of the above-named lines is given in Fig. 66, from which we

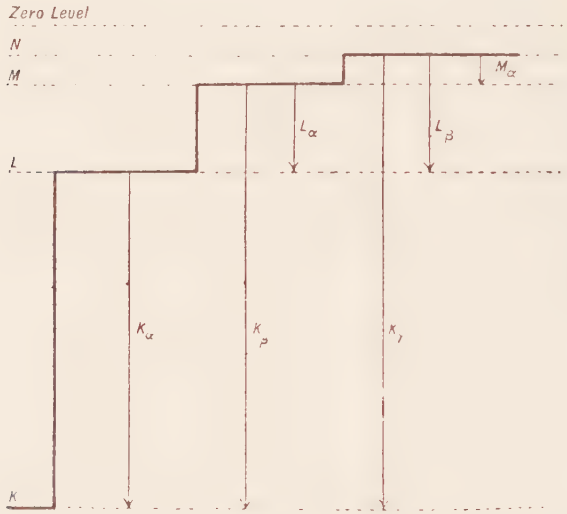



FIG. 66.—Diagrammatic scheme of the most important X-ray lines.

obtain at once the following approximate *combination relations* for the *frequencies* :—

$$(3) \qquad \begin{cases} K_\beta = K_\alpha + L_\alpha, \\ K_\gamma = K_\alpha + L_\beta = K_\beta + M_\alpha, \\ L_\beta = L_\alpha + M_\alpha. \end{cases}$$

Soon after the discovery of X-ray spectra, experiment led to the important rule that the *excitation potential* necessary for the *generation* of a definite line is always *equal to the characteristic potential of the final level*, the potential associated with any term being governed by § 93 eqn. (7). The converse also holds good. If the potential be gradually increased, then, on reaching a certain value, *all the lines* for which this potential represents the *common final level* will be *simultaneously excited*.

From this fact *Kossel*, in 1914, first drew the important conclusion that the *emission of an X-ray line* can only occur when an electron in the emitting atom has been *previously* removed from its normal orbit to *beyond the limits of the atom*. Thus, a *place in the atom* must first, as it were, *fall vacant*. If this vacant place corresponds to a quantity of energy $(-h\nu'')$, and if it be filled by an electron which was previously situated at an energy-level $(-h\nu')$, a line of frequency $(\nu'' - \nu')$ will appear. For the generation of this line, however, the *energy* $h\nu''$ must have been *previously expended*.

A convincing proof of the accuracy of Kossel's conception is afforded by the appearance of *absorption band heads in continuous X-ray spectra*. If the *X-ray line spectrum* of an element be *excited by primary continuous X-ray radiation*, instead of by cathode rays, the *wave-lengths characteristic of the energy-levels in the line spectrum excited* manifest themselves as heads of absorption bands in the primary X-ray spectrum.

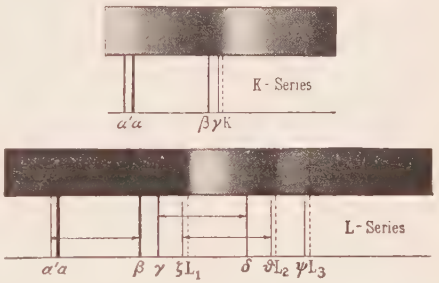



FIG. 67.—Absorption band heads in the continuous X-ray spectrum.

The explanation of this phenomenon follows immediately from Kossel's conception. For, in accordance with Einstein's law of light-quanta, we must suppose *light-quanta* to be *absorbed from the primary continuous radiation* and to be *transformed into light-quanta of the secondary line radiation*. Hence, according to Kossel's theory, a line in the *K*-series can be excited only by primary rays of frequency *at least* as great as the characteristic frequency of the one-quantum level, the so-called *K*-level. On the other hand, in the excitation of *K*-lines, no primary rays are absorbed, the wave-lengths of which are larger than the wave-length characterizing the *K*-level. Thus, we may expect to find the head of an absorption band at this wave-length, and likewise at the wave-lengths characteristic of the two- and three-quantum *L*- and *M*-levels.

These expectations are borne out by actual experiment. In the upper photograph reproduced in Fig. 67, we see the absorption band head of the *K*-series in silver; beneath it

are inserted the four lines of the corresponding *K*-series (their wave-lengths being larger). The primary X-radiation was generated in a 40 kilovolt tube, so that, according to § 93 eqn. (3), the continuous X-ray spectrum extended down to about 300 X. The sharp head to the absorption band, seen in the photograph, corresponds to a wave-length of 485 X. Since the wave-lengths of the absorption band heads are characteristic of the energy-levels themselves, thus corresponding to the difference between a level and the zero-line, the *heads of the absorption bands* also represent *series limits*, as is likewise seen from Fig. 67.

The problem of experimental X-ray spectroscopy consists chiefly in the *exact determinations of wave-lengths*, both of absorption band heads and of individual emission lines. Much the most accurate method of measurement is the *direct* one based on the laws *of crystal reflection*, which we shall discuss later; but this method can only be applied within a limited region of wave-lengths. The majority of wave-lengths are measured by means of a *rock-salt* or *calcite crystal*, which can be used for wave-lengths up to about 5,000 X. Wave-lengths more than twice as large can be reached with the help of *selenite* or *mica*, which have a larger lattice constant. Exact measurements by means of crystals have been perfected, principally by *Siegbahn*, to such a degree that, at the present time, X-ray wave-lengths can be determined with an accuracy of about a millionth part of the measured value. (That is to say, wave-lengths can be compared with one another, or referred to an idealized lattice constant with that accuracy.) [2]

In comparison with these direct precision methods, only a secondary value can be attached to the *indirect methods*. These are, however, of the greatest importance in the *boundary region* between the largest wave-lengths measurable with crystals and the shortest wave-lengths in the ultra-violet. The most important of these indirect methods consists in the derivation of the wave-length of the *series limits* from the measured excitation potentials (defined by the values at which absorption suddenly commences). In exceptional cases, the characteristic wave-lengths can also be measured with an *optical diffraction grating* when the *nuclear charge numbers* are *small*, because, with decreasing nuclear charge number, all series gradually shift over to the ultra-violet region.

[2] *Vide* § 120.

The number of lines belonging to a given series is, in general, fairly large: thus, as may be seen from Fig. 68, the *L*-series of rhodium contains eleven lines.[3] We must, therefore, assume at any rate *several* *N*-, *M*-, and *L*-levels. This necessity is also rendered evident by researches on absorption band heads, which show that the *L*-limit is *three-fold* and the *M*-limit *five-fold* in those elements where they are observed. Only the *K*-limit appears to be invariably single (*cf.* the previous Fig. 67).

The *scheme of X-ray terms*, which was established in 1922 by *Smekal*, *Wentzel*, *Coster*, and *Bohr*, exhibits a far-reaching similarity to the term system of the *alkali metals*. This

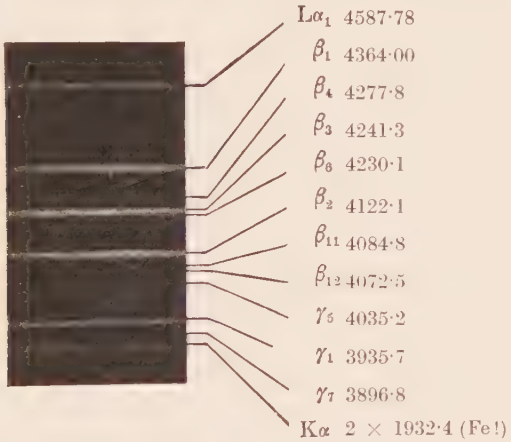



FIG. 68.—*L*-series of rhodium (after Siegbahn).

may be readily understood, for, just as the alkali spectra originate in the transitions of electrons, so do X-ray spectra arise from "*jumping a gap*" between any complete and another, previously incomplete, sub-group. We find in the system of X-ray terms also a *doublet system*, in which *s*-, *p*-, *d*-, and *f*- terms can be distinguished. The one-quantum terms are called *K*-terms, the two-quantum *L*-terms, the three-, four-, five, and six-quantum terms *M*-, *N*-, *O*-, and *P*-terms respectively.

[3] The lines have been marked with the corresponding wave-lengths in X. As well as the rhodium lines, the $K_\alpha$-line of iron is also to be seen in Fig. 68, iron being present as a small impurity in the rhodium. In X-ray spectroscopy it is possible to detect impurities amounting to only from a hundredth to a tenth of 1 *per cent.*

In Fig. 69 is reproduced the *scheme of X-ray terms* as determined for *elements of the highest atomic numbers* (from radon onwards). With these highest elements there are 1 $K$-term, 3 $L$-terms, 5 $M$-terms, 7 $N$-terms, 5 $O$-terms, and 3 $P$-terms. In the figure single vertical dashes on the left-hand side denote those terms which gradually lapse as we descend the natural sequence of the elements from radon

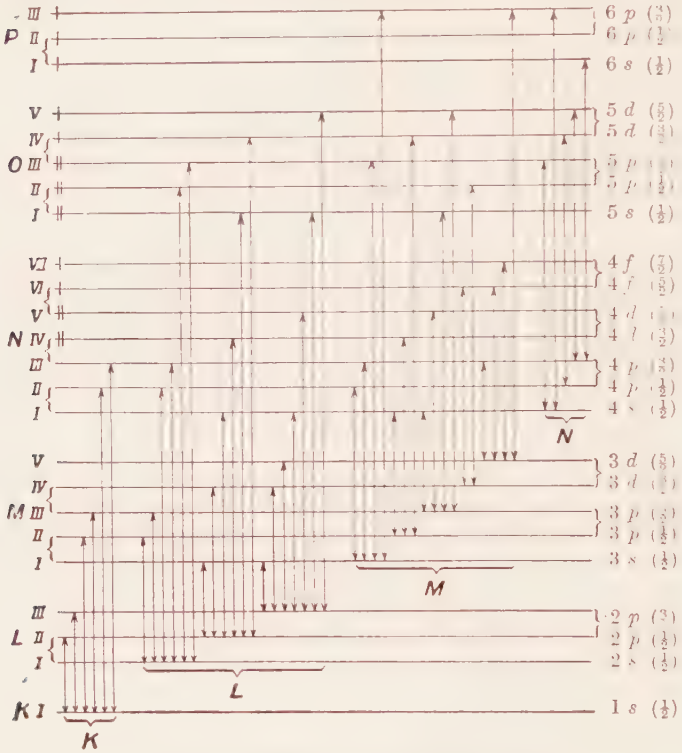



Fig. 69.—Scheme of X-ray terms and lines.

to xenon. In the case of the elements near krypton, other terms are also lacking, these being denoted in the figure by double vertical dashes. On the right-hand side of Fig. 69 are shown the symbols for the X-ray terms in the terminology of optical spectra: the principal quantum number is given first, followed by the customary symbol for the subordinate quantum number ($s$ for 0, $p$ for 1, $d$ for 2, and $f$ for 3) and, finally, in brackets, the inner quantum number

($j$). It follows from § 100 eqns. (7) and (13) that *permissible combinations* between the terms are governed by the relations

$$\Delta l = \pm 1, \quad \Delta j = \pm 1 \text{ or } 0. \tag{4}$$

The combinations obtained in this way are shown in Fig. 69 ; they are confirmed by experiment. For example, the transition from an $L$- to a $K$-level is associated with two lines, while seven lines correspond to the transition from an $M$- to an $L$-level. It is also evident from Fig. 69 how the individual groups of lines are resolved into sub-groups distinguished by the same final level, and thus by the same absorption band head.

The X-ray terms are known at the present time with fair accuracy for most elements. They were determined either directly by measurements of absorption band heads or, with the help of the above-given scheme of lines, by measurements of the emission lines. The relations between the terms and the atomic numbers are shown in Fig. 70, where the ordinates represent the *square roots* of the terms referred to the Rydberg term as unit.[4] The lines drawn underneath the figure will be considered later (§ 112). The course of the various curves is more or less linear, in accordance with Moseley's Law ; the larger the principal quantum number, the smaller is the angle between the curve and the axis of abscissæ.

Rules similar to those governing alkali doublets apply also to the *doublets of X-ray terms*. Terms with the same principal and subordinate quantum number, and differing only in their *inner* quantum number, form so-called *regular doublets :* these are marked in Fig. 69 by brackets on the *right-hand* side. Just as with the alkali terms, the separation of regular doublets in X-ray terms increases rapidly with the atomic number. As *Sommerfeld* discovered in 1916, the frequency difference of a *regular* doublet is given by the simple relation

$$\Delta\nu = \Delta\nu_H\,(z - b)^4, \tag{5}$$

where $\Delta\nu_H$ denotes the so-called frequency difference of the hydrogen doublet [§ 104 eqn. (26) ], $z$ the atomic number, and $b$ the relevant screening constant.

[4] The terms are therefore obtained on multiplying the squares of the ordinates by 109,737. The $P$-terms, which are still rather uncertain and have, of course, only been calculated for the highest of the elements, have been omitted in Fig. 70.




FIG. 70.—X-ray terms of the elements.

Wave-length difference and frequency difference are connected by the simple relation

$$\Delta\lambda = -\frac{c}{\nu^2}\Delta\nu, \tag{6}$$

which is at once obtained by differentiating the equation

$$c = \lambda\nu.$$

Since, by Moseley's law, the square root of the frequency is proportional to the atomic number less the screening constant, it follows from eqns. (5) and (6) that the *wave-length difference of a regular doublet* is almost *independent of the atomic number*. For the rest, Fig. 70 clearly shows that, just as with the alkali terms, so also with the X-ray terms, the regular doublet separation diminishes as the principal and the subordinate quantum number increase.

Adjacent terms with the same inner quantum number, but with a different subordinate quantum number, constitute so-called *irregular doublets*;[5] these are indicated in Fig. 69 by brackets on the *left-hand* side. Their frequency difference increases very slowly with the atomic number, as was discovered in 1920 by *G. Hertz*. Hence, in Fig. 70, the curves of two terms forming an *irregular doublet* have a nearly *parallel* course, whereas the curves of two terms constituting a *regular doublet diverge* more and more as the atomic number increases.

Furthermore, it may be noted that, with some elements, besides the lines indicated by the scheme of Fig. 69, a few other, weaker, lines have been detected. These are evidently produced by the removal of *more than a single* electron from the interior of the atom, and they belong to the so-called *X-ray spark spectra*, which were first investigated by *Wentzel*.[6] It should also be mentioned that *X-ray terms* have been found to *depend slightly on the molecular constitution*.[7] Repeated attempts have been made to discover *absorption-lines* in X-ray spectra, but without success. The complete absence of such lines represents the essential difference between X-ray and optical spectra.

[5] Formerly, regular doublets were termed "relativistic," and irregular "screening doublets." An exact derivation of the doublet formulæ on the basis of quantum mechanics was first effected by Heisenberg, Jordan, Pauli, and Dirac.

[6] G. Wentzel, *Zeitschr. f. Phys.*, **31** (1925), p. 445.

[7] For this and other details of the progress of X-ray spectroscopy between 1922 and 1925, *vide* the summary by A. E. Lindh in the *Physikal. Zeitschr.*, **28** (1927), pp. 24 and 93.

## § 112. The Periodic System.

The important conclusions concerning atomic structure and the system of the elements, to which optical spectra lead, are entirely confirmed by X-ray spectra. In point of fact, the *scheme of X-ray terms* (*cf.* the short vertical single and double strokes at the left-hand side of Fig. 69) clearly shows that, at the end of the fourth period (*i.e.*, with krypton), the four-quantum *d*- and *f*-electrons have not yet appeared. We see also that the fifth period is completed without 4 *f*- and 5 *d*-electrons, and that, even in the highest elements, the fifth-quantum group contains only *s*-, *p*-, and *d*-electrons, and the sixth quantum group only *s*- and *p*-electrons.

A particularly strong confirmation of the conclusions as to the *completion of the inner groups of electrons*, drawn from optical spectra, is afforded by the curves representing the *dependence of the square roots of the X-ray spectroscopic terms of the elements on the atomic members* (*cf.* the previous Fig. 70). In accordance with Moseley's law, the form of these curves should be exactly *rectilinear*, provided that an increase in the atomic number by unity is accompanied by a like increase in the *effective fraction of the nuclear charge*, for the term concerned, and a corresponding increase in the *binding energy*. During the development of an inner electron group, however, the increase in the binding energy and hence of the term must at any rate be interrupted or, at the very least, retarded. This is due to the fact that the increase in the nuclear charge is entirely or partially counterbalanced by a similar increase of one in the number of *inner electrons screening* the nucleus. Hence, *during the development of an inner group of electrons* of given quantum number, the *curves* for the terms of *higher quantum number* should run *horizontally*, or at least at a *smaller angle of inclination* to the axis of abscissæ.

That such is the case is shown by Fig. 70 (*cf.* the lines drawn beneath the figure), and still more distinctly by Fig. 71; the latter refers to the elements between Nos. 50 and 74, and is by *Nishina*,[1] to whom very accurate new determinations of the X-ray terms are due.[1] In this figure the horizontal course of the five-quantum *O*-terms during the completion of the four-quantum group from 18 to 32 electrons is clearly visible. Moreover, during this completion, the curves for the *N*-terms run less steeply, for the addition of new electrons in the same group has, at least, a partial

[1] Y. Nishina, *Phil. Mag.*, **49** (1925), p. 521.

screening effect. The completion of the 4 *f*-group with the last of the rare earths, element No. 71, is expressed by the sudden change in the direction of the curves for $N_{VI}$ and $N_{VII}$ at that number. From the previous Fig. 70 it is also manifest that the 4 *f*-term is at first smaller than the 5 *s*-

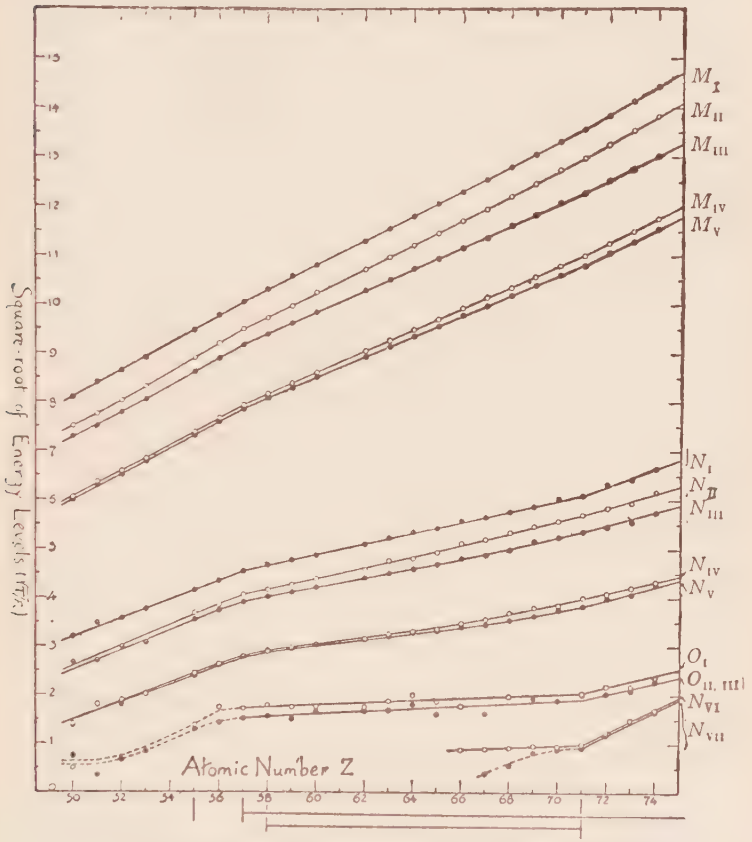



FIG. 71.—X-ray terms of elements Nos. 50–74 (after Nishina).

and 5 *p*-terms, and at the very beginning is even smaller than the 5 *d*-term. As soon as the development of the four-quantum group of electrons is completed [at 71 (Lu)], however, the binding energy of the 4 *f*-term rapidly increases, so that, between the 75th and the 83rd element, the 4 *f*-curves ***cross*** **the 5 *s*- and 5 *p*-curves.**

Since, in the elements of the sixth period, the four-quantum group lies fairly *deep within the atom*, during its completion, the *chemical* and also those *physical properties* such as valency and optical-spectral behaviour, which are conditioned by the *peripheral electrons*, must remain practically *unchanged*. In point of fact, the elements following after lanthanum closely resemble one another and lanthanum in chemical and in important physical properties, whereas lanthanum is very dissimilar from the immediately preceding element, barium. This fact confirms that the completion of the four-quantum group from 18 to 32 electrons *begins* with the element next after lanthanum, *cerium* (58). Corresponding to the difference between the numbers 32 and 18, there must naturally be a total of 14 elements resembling lanthanum, or, including this element, *fifteen rare earths* altogether.

The agreement of the peripheral atomic properties does, in fact, extend over 15 successive elements, beginning with lanthanum and ending with element 71, *lutecium*. At the time (1922) when *Bohr* established his theory of the periodic system, element 72 was not yet known. Up till then it had commonly been supposed that it must also be a rare earth, and searches were therefore made for it in those minerals in which other rare earths had been found. It followed from the quantum theory of the system of the elements, however, that the unknown element could not possibly be a rare earth (because the theory does not permit of more than 15 rare earths), but rather that it must be a *homologue of zirconium*, which occupies the fourth place in the fifth period. Since homologous elements frequently occur together in nature, the idea suggested itself of looking for the unknown element in zirconium minerals. It constituted a brilliant triumph for Bohr's new theory when, in 1922, *Hevesy* and *Coster* actually succeeded in discovering the new element in zirconium minerals; in honour of the city of Copenhagen, where the discovery was made, they named it "*hafnium*." In point of fact, hafnium is by no means rare; there is, for example, only about five times more nickel than hafnium in the earth's crust. It can only be ascribed to its close resemblance to zirconium that hafnium was not discovered long ago, and also that it has not yet been possible to prepare it in a pure state.[2]

[2] The discovery was made by X-ray spectroscopy. *Cf.* the summary by Paneth in Vol. II. of "*Ergebnisse der exakten Naturwissenschaften*" (Berlin; Springer, 1923); also G. v. Hevesy, "*Das Element Hafnium*" (Berlin; Springer, 1927).

In the *seventh period* of the system of the elements, the six-quantum group must be completed from 8 to 18, and, in addition, the five-quantum group from 18 to 32 electrons. From the homology of actinium and lanthanum, of protoactinium and tantalum, and of uranium and tungsten, we must conclude that the completion of the second outermost group of electrons (in this case, the six-quantum group) also begins in the seventh period with the third element, actinium, just as in the fourth, fifth, and sixth periods. On the other

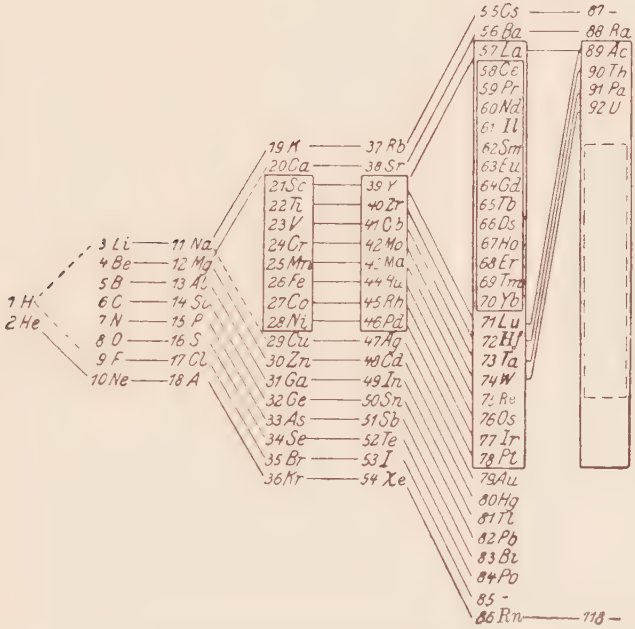



FIG. 72.—The periodic system.

hand, up to and including uranium, the number of electrons in the five-quantum group further within the atom does not increase.

The arrangement of the periodic system reproduced in Fig. 72 clearly distinguishes from among the whole series of elements those in which an *inner group of electrons* contains *incomplete sub-groups*. In point of fact, the elements enclosed in the rectangles in Fig 72 exhibit, with their incomplete inner atomic structure, peculiarities lack-

TABLE XV

*Grouping of Electrons in the Periodic System*

| | 1 *s* | 2 *s p* | 3 *s p d* | 4 *s p d f* | 5 *s p d* | 6 *s p d* | 7 *s* |
|---|---|---|---|---|---|---|---|
| 1 H | 1 | | | | | | |
| 2 He | 2 | | | | | | |
| 3 Li | 2 | 1 | | | | | |
| 4 Be | 2 | 2 | | | | | |
| 5 B | 2 | 2 1 | | | | | |
| 6 C | 2 | 2 2 | | | | | |
| 7 N | 2 | 2 3 | | | | | |
| 8 O | 2 | 2 4 | | | | | |
| 9 F | 2 | 2 5 | | | | | |
| 10 Ne | 2 | 2 6 | | | | | |
| 11 Na | 2 | 2 6 | 1 | | | | |
| 12 Mg | 2 | 2 6 | 2 | | | | |
| 13 Al | 2 | 2 6 | 2 1 | | | | |
| 14 Si | 2 | 2 6 | 2 2 | | | | |
| 15 P | 2 | 2 6 | 2 3 | | | | |
| 16 S | 2 | 2 6 | 2 4 | | | | |
| 17 Cl | 2 | 2 6 | 2 5 | | | | |
| 18 A | 2 | 2 6 | 2 6 | | | | |
| 19 K | 2 | 2 6 | 2 6 | 1 | | | |
| 20 Ca | 2 | 2 6 | 2 6 | 2 | | | |
| 21 Sc | 2 | 2 6 | 2 6 1 | 2 | | | |
| 22 Ti | 2 | 2 6 | 2 6 2 | 2 | | | |
| 23 V | 2 | 2 6 | 2 6 3 | 2 | | | |
| 24 Cr | 2 | 2 6 | 2 6 5 | 1 | | | |
| 25 Mn | 2 | 2 6 | 2 6 5 | 2 | | | |
| 26 Fe | 2 | 2 6 | 2 6 6 | 2 | | | |
| 27 Co | 2 | 2 6 | 2 6 7 | 2 | | | |
| 28 Ni | 2 | 2 6 | 2 6 8 | 2 | | | |
| 29 Cu | 2 | 2 6 | 2 6 10 | 1 | | | |
| 30 Zn | 2 | 2 6 | 2 6 10 | 2 | | | |
| 31 Ga | 2 | 2 6 | 2 6 10 | 2 1 | | | |
| 32 Ge | 2 | 2 6 | 2 6 10 | 2 2 | | | |
| 33 As | 2 | 2 6 | 2 6 10 | 2 3 | | | |
| 34 Se | 2 | 2 6 | 2 6 10 | 2 4 | | | |
| 35 Br | 2 | 2 6 | 2 6 10 | 2 5 | | | |
| 36 Kr | 2 | 2 6 | 2 6 10 | 2 6 | | | |
| 37 Rb | 2 | 2 6 | 2 6 10 | 2 6 | 1 | | |
| 38 Sr | 2 | 2 6 | 2 6 10 | 2 6 | 2 | | |
| 39 Y | 2 | 2 6 | 2 6 10 | 2 6 1 | 2 | | |

TABLE XV—continued

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| | *s* | *s p* | *s p d* | *s p d f* | *s p d* | *s p d* | *s* |
| 40 Zr | 2 | 2 6 | 2 6 10 | 2 6 2 | 2 | | |
| 41 Cb | 2 | 2 6 | 2 6 10 | 2 6 4 | 1 | | |
| 42 Mo | 2 | 2 6 | 2 6 10 | 2 6 5 | 1 | | |
| 43 Ma | 2 | 2 6 | 2 6 10 | 2 6 (5) | (2) | | |
| 44 Ru | 2 | 2 6 | 2 6 10 | 2 6 7 | 1 | | |
| 45 Rh | 2 | 2 6 | 2 6 10 | 2 6 8 | 1 | | |
| 46 Pd | 2 | 2 6 | 2 6 10 | 2 6 10 | | | |
| 47 Ag | 2 | 2 6 | 2 6 10 | 2 6 10 | 1 | | |
| 48 Cd | 2 | 2 6 | 2 6 10 | 2 6 10 | 2 | | |
| 49 In | 2 | 2 6 | 2 6 10 | 2 6 10 | 2 1 | | |
| 50 Sn | 2 | 2 6 | 2 6 10 | 2 6 10 | 2 2 | | |
| 51 Sb | 2 | 2 6 | 2 6 10 | 2 6 10 | 2 3 | | |
| 52 Te | 2 | 2 6 | 2 6 10 | 2 6 10 | 2 4 | | |
| 53 I | 2 | 2 6 | 2 6 10 | 2 6 10 | 2 5 | | |
| 54 Xe | 2 | 2 6 | 2 6 10 | 2 6 10 | 2 6 | | |
| 55 Cs | 2 | 2 6 | 2 6 10 | 2 6 10 | 2 6 | 1 | |
| 56 Ba | 2 | 2 6 | 2 6 10 | 2 6 10 | 2 6 | 2 | |
| 57 La | 2 | 2 6 | 2 6 10 | 2 6 10 | 2 6 1 | 2 | |
| 58 Ce | 2 | 2 6 | 2 6 10 | 2 6 10 (1) | 2 6 (1) | (2) | |
| . . | . | . . | . . . | . . . . | . . . | . | |
| . . | . | . . | . . . | . . . . | . . . | . | |
| 72 Hf | 2 | 2 6 | 2 6 10 | 2 6 10 14 | 2 6 (2) | (2) | |
| 73 Ta | 2 | 2 6 | 2 6 10 | 2 6 10 14 | 2 6 (3) | (2) | |
| 74 W | 2 | 2 6 | 2 6 10 | 2 6 10 14 | 2 6 4 | 2 | |
| 75 Re | 2 | 2 6 | 2 6 10 | 2 6 10 14 | 2 6 (5) | (2) | |
| 76 Os | 2 | 2 6 | 2 6 10 | 2 6 10 14 | 2 6 (6) | (2) | |
| 77 Ir | 2 | 2 6 | 2 6 10 | 2 6 10 14 | 2 6 (7) | (2) | |
| 78 Pt | 2 | 2 6 | 2 6 10 | 2 6 10 14 | 2 6 8 | 2 | |
| 79 Au | 2 | 2 6 | 2 6 10 | 2 6 10 14 | 2 6 10 | 1 | |
| 80 Hg | 2 | 2 6 | 2 6 10 | 2 6 10 14 | 2 6 10 | 2 | |
| 81 Tl | 2 | 2 6 | 2 6 10 | 2 6 10 14 | 2 6 10 | 2 1 | |
| 82 Pb | 2 | 2 6 | 2 6 10 | 2 6 10 14 | 2 6 10 | 2 2 | |
| 83 Bi | 2 | 2 6 | 2 6 10 | 2 6 10 14 | 2 6 10 | 2 3 | |
| 84 Po | 2 | 2 6 | 2 6 10 | 2 6 10 14 | 2 6 10 | 2 4 | |
| 85 — | 2 | 2 6 | 2 6 10 | 2 6 10 14 | 2 6 10 | 2 5 | |
| 86 Rn | 2 | 2 6 | 2 6 10 | 2 6 10 14 | 2 6 10 | 2 6 | |
| 87 — | 2 | 2 6 | 2 6 10 | 2 6 10 14 | 2 6 10 | 2 6 | 1 |
| 88 Ra | 2 | 2 6 | 2 6 10 | 2 6 10 14 | 2 6 10 | 2 6 | 2 |
| 89 Ac | 2 | 2 6 | 2 6 10 | 2 6 10 14 | 2 6 10 | 2 6 (1) | (2) |
| 90 Th | 2 | 2 6 | 2 6 10 | 2 6 10 14 | 2 6 10 | 2 6 (2) | (2) |
| 91 Pa | 2 | 2 6 | 2 6 10 | 2 6 10 14 | 2 6 10 | 2 6 (3) | (2) |
| 92 U | 2 | 2 6 | 2 6 10 | 2 6 10 14 | 2 6 10 | 2 6 (5) | (1) |

ing in the other elements. They are frequently *multivalent*; they form *coloured salts*; and, as a rule, they are *paramagnetic*. The multivalency is evidently due to the possibility that electrons belonging to an incompletely developed inner group may be just as *loosely* bound as the peripheral valency electrons.[3] Hence, like the latter, they are also able to absorb quanta of *visible* light, so that their ions are coloured, despite the absence of valency electrons. Finally, as we shall discuss more fully in a subsequent section (§ 116), paramagnetism is ascribed to an incomplete balance in the magnetic moments of the electrons in the individual atoms, and thus likewise appears as a consequence of asymmetry in the internal structure of the atom.

As we have already stated, the *quantum theory of the system of the elements* was established in 1922 by Bohr. The essential assumption in this theory of the development of inner groups of electrons had, indeed, been anticipated by *Ladenburg*, who sought to refer the chemical and magnetic characteristics of the iron, palladium, and platinum series to the formation of an "*intermediate shell*" of loosely bound electrons. In 1922 Bohr distinguished, quite correctly, within each principal group the same number of sub-groups as the principal quantum number. At that time, however, he assumed, erroneously, that in a complete group the electrons are distributed equally among the sub-groups, so that, for example, the complete three-quantum group was supposed to consist of three sub-groups, each of six electrons. It was first demonstrated by *Stoner*[4] and *Main Smith*[5] in 1924 that we must assume the numbers of electrons in the complete sub-groups to be 2, 6, 10, and 14. The rules established by these two investigators were first explained by *Pauli's principle*, which, in connection with the experimentally derived *fundamental terms*, led to a determination of the electronic distribution for most of the elements, as has been described in the previous sections. With the addition of the elements in the seventh period, but omitting the rare earths, the grouping of the electrons in the periodic

[3] We have, moreover, to assume such a loose binding of the inner electrons in the elements which, though lying outside the rectangles in Fig. 72, yet border on them. Thus *copper*, as already mentioned, is both univalent in *cuprous* compounds and divalent in *cupric* compounds. The distribution $s\,d^{10}$ corresponds to the former, and $s^2\,d^9$ to the latter. Actually, cuprous ions are colourless and cupric ions coloured.

[4] E. C. Stoner, *Phil. Mag.*, **48** (1924), p. 719.

[5] J. D. Main Smith, *Journ. Soc. Chem. Ind.*, **44** (1925), p. 944.

system, as given by Hund [6] and developed by *McLennan*,[7] is reproduced in Table XV.

## § 113. The Zeeman Effect.

The *resolution of spectral lines in a magnetic field* was discovered by *Zeeman* in 1896. *Lorentz* immediately developed a theory for the phenomenon, and this theory permitted of the first calculation of the mass of an electron. Further investigation of the so-called *Zeeman effect* led to a distinction between two varieties of the phenomenon, which are known as the *normal* and the *anomalous effect*. In the *normal* Zeeman effect, besides the original line (which, however, in certain circumstances to be explained subsequently, may be invisible), *two other lines* are to be observed: these lie at an *equal distance* on either side of the original line, the frequency difference for each being determined by the relation

$$\frac{\Delta\nu}{c} = \frac{\omega}{c} H. \tag{1}$$

In this expression, $H$ denotes the magnetic strength of field, and $\omega$ the universal constant defined in § 102 eqn. (8), *viz.*, $\epsilon 4\pi\mu c$. Hence, if the field-strength be stated in absolute units—"gauss"—

$$\frac{\Delta\nu}{c} = 4{\cdot}70 \times 10^{-5} H. \tag{2}$$

With fields of a few ten thousand gauss (in the present state of experimental physics a permanent field of more than about 100,000 gauss cannot be realized), the separation of the Zeeman effect is of the order of 1 Å. in the visible region of the spectrum. In the *anomalous* effect, the resolution of the spectral lines is far more complicated in character than in the normal effect, but here too a series of rules has been empirically derived.

The quantum theory of the atom attributes the difference between the normal and the anomalous effect to a distinction between *weak* and *strong magnetic fields*. As already mentioned, this distinction has no absolute meaning; a field is weak or strong with respect to a given spectral term according as the resolution into multiplets preponderates over the magnetic separation or contrariwise. In this sense, the *normal* Zeeman-effect appears as the conse-

[6] F. Hund, *l.c.*

[7] J. C. McLennan, A. B. McLay, and H. G. Smith, *Proc. Roy. Soc. A.*, **112** (1926), p. 76.

quence of a *strong* magnetic field. As was previously demonstrated—§ 102 eqn. (21)—a term $T_1$ is transformed in a strong field according to the relation

$$T_1' = T_1 + m_1 \frac{\omega}{c} H, \tag{3}$$

where $m_1$ denotes the magnetic quantum number. Similarly, for a second term $T_2$ we have

$$T_2' = T_2 + m_2 \frac{\omega}{c} H. \tag{4}$$

Let $\nu$ be the frequency corresponding to the combination of the terms $T_1$ and $T_2$ in the absence of a magnetic field, and $\nu'$ the frequency when a field is present. Since, however, in accordance with § 102 eqn. (3), the magnetic quantum number can only alter by unity or zero on a change of atomic state, we must have either

$$\nu' = \nu \pm \omega H \tag{5}$$

or

$$\nu' = \nu. \tag{6}$$

In the *normal* Zeeman effect, therefore, the original line is replaced in general by an *equidistant line triplet*, the central line of which *coincides* with the original one. If, now, the magnetic quantum number changes by unity, it follows from the definition of $m$ that the *angular momentum* of the atom emitting the light alters in the direction of the field. According to the Law of the *Conservation of Angular Momentum*, we must assume that the alteration appears as an imparted angular momentum to the *emitted light-quantum*; in other words, that *circularly polarized light* is *emitted*.[1] On the other hand, when $m$ remains *unchanged*, we must assume light to be emitted by *linear* vibrations of the vector of the electric field-strength, and we may well suppose the oscillations to lie *in the direction of the field*, just as, in the other case, the circular vibrations take place in a *plane at right angles* to the field.

Now it is known from classical optics that light vibrations are propagated only in directions which are normal to the direction of vibration. Hence, in the so-called *longitudinal effect*, when observation is made in the direction of the field, the central line will be missing, so that only a

[1] This important result was first obtained in 1918 by *Rubinowicz* as a general so-called "*polarization rule*," together with the "*selection rule*" which states that the quantum number determining the angular momentum of the atom can only vary by $\pm$ 1 or 0.

*line doublet* will be visible, the two components being *circularly polarized* in opposite senses.[2] In the *transversal effect*, the whole triplet must be visible, but with *linearly polarized lines* throughout, the vibrations for the two outer lines being at right angles to the direction of field, whereas those for the central line are parallel to the latter. In point of fact, observation confirms all these theoretical results (*cf.* Fig. 73).

The phenomena of the *anomalous* Zeeman effect can be correctly predicted with the help of Landé's resolution formula, derived in a previous section (§ 102). This formula also affords an explanation for two rules which had been obtained empirically long before and are known as *Preston's* and *Runge's rules*. As Preston discovered in 1899, *cognate lines*, *i.e.*, lines representing combinations between homogeneous terms, exhibit the same anomalous effect which, moreover, is found to be independent of the chemical nature of the atom emitting them. By homogeneous terms we mean, in the sense of the quantum theory, those which have the same multiplicity and character—the subordinate quantum number as expressed by the symbols *S*, *P*, *D*, *F*, *etc.* It follows from Preston's rule that the resolution of the different lines in a series will be the same, and, furthermore, that we shall find the same resolution for two lines corresponding to opposite combinations. That is to say, if one series arises from the combination of terms of type I with a certain term of type II, and a second series is due to the combination of terms of type II with a certain term of type I, the lines in both series will be similarly resolved. Actually, for example, the lines in the *principal series* of the alkali metals exhibit the *same resolution* as the lines in the *sharp subordinate series*.

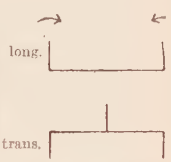


FIG. 73.—Scheme of the normal Zeeman effect.

Preston's rule, which forms a valuable guide for the term analysis of optical spectra, is naturally a necessary consequence of Landé's formula, according to which the resolving factor is independent both of the principal quantum number and of all chemical characteristics, and is determined by the subordinate quantum number and the multiplicity. Landé's formula also yields the rule established in 1907 by

[2] It may be shown with the help of the classical theory that, in the case of the *component with the smaller wave-length*, the sense of the circular polarization will be that of the positive current in the coils of the electro-magnet which creates the magnetic field.

*Runge.* This states that the separations in the anomalous Zeeman effect are always *rational multiples of the normal separation, i.e.*, that, referred to the latter as unit, they are of the form $z/u$, where both $z$ (the Runge numerator) and $u$ (the Runge denominator) are *integers.*

In the case of the *alkali metals* and *alkaline earths*, the anomalous Zeeman effect is comparatively simple. The corresponding $g$-values, as derived from Landé's formula and also empirically obtained, are contained in Table XVI.

TABLE XVI

*Resolving Factors*

| | Alkali metals. | | | | Alkaline earths. | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Doublet terms. | | | | Singlet terms. | | | | Triplet terms. | | | | |
| $j$-values. | $\frac{1}{2}$ | $\frac{3}{2}$ | $\frac{5}{2}$ | $\frac{7}{2}$ | 0 | 1 | 2 | 3 | 0 | 1 | 2 | 3 | 4 |
| $S$ | 2 | | | | $\frac{0}{0}$ | | | | | 2 | | | |
| $P$ | $\frac{2}{3}$ | $\frac{4}{3}$ | | | | 1 | | | $\frac{0}{0}$ | $\frac{3}{2}$ | $\frac{3}{2}$ | | |
| $D$ | | $\frac{4}{5}$ | $\frac{6}{5}$ | | | | 1 | | | $\frac{1}{2}$ | $\frac{7}{6}$ | $\frac{4}{3}$ | |
| $F$ | | | $\frac{6}{7}$ | $\frac{8}{7}$ | | | | 1 | | | $\frac{2}{3}$ | $\frac{13}{12}$ | $\frac{5}{4}$ |

For $j = 0$ no separation occurs, corresponding to the *indeterminate* value $0/0$ for the resolving factor. Furthermore, it appears from Table XVI that lines arising from

TABLE XVII

*Variations of the Alkali Terms in the Anomalous Zeeman Effect*

| $m =$ | $-\frac{5}{2}$ | $-\frac{3}{2}$ | $-\frac{1}{2}$ | $+\frac{1}{2}$ | $+\frac{3}{2}$ | $+\frac{5}{2}$ |
|---|---|---|---|---|---|---|
| $S$ | | | $-1$ | $+1$ | | |
| $P_{\frac{1}{2}}$ | | | $-\frac{1}{3}$ | $+\frac{1}{3}$ | | |
| $P_{\frac{3}{2}}$ | | $-\frac{6}{3}$ | $-\frac{2}{3}$ | $+\frac{2}{3}$ | $+\frac{6}{3}$ | |
| $D_{\frac{3}{2}}$ | | $-\frac{6}{5}$ | $-\frac{2}{5}$ | $+\frac{2}{5}$ | $+\frac{6}{5}$ | |
| $D_{\frac{5}{2}}$ | $-\frac{15}{5}$ | $-\frac{9}{5}$ | $-\frac{3}{5}$ | $+\frac{3}{5}$ | $+\frac{9}{5}$ | $+\frac{15}{5}$ |

combinations of two singlet terms must yield the normal triplet.

If we consider first the *alkali metals*, we obtain the values reproduced in Table XVII for the *variations* from the term in the absence of a field, the normal separation being chosen as unit. Let us take as an example of this scheme the well-known *yellow sodium line*, the so-called *D*-line.[3] This line is the first line of the principal series and consists of the two components

$$D_1 = 5895{\cdot}93 \text{ Å}. \ (^2P_{\frac{1}{2}} \rightarrow {}^2S),$$
$$D_2 = 5889{\cdot}96 \text{ Å}. \ (^2P_{\frac{3}{2}} \rightarrow {}^2S).$$

Bearing in mind the combination rule, according to which $\Delta m$ must be equal to $\pm 1$ or 0, we obtain for the line variations, which represent the *differences* of the term variations, the following scheme, where underlined values correspond to the condition $\Delta m = 0$ :—

$$(7) \quad \begin{cases} D_1\text{-line}, & & -\frac{4}{3}, & & \underline{-\frac{2}{3}}, \ \underline{+\frac{2}{3}}, & & +\frac{4}{3}, \\ D_2\text{-line}, & -\frac{5}{3}, & & -\frac{3}{3}, \ \underline{-\frac{1}{3}}, \ \underline{+\frac{1}{3}}, & +\frac{3}{3} & & +\frac{5}{3}. \end{cases}$$

The $D_1$-line therefore undergoes a *four-fold*, and the $D_2$-line a *six-fold* resolution, so that, in a weak magnetic field, the original doublet is replaced by a multiplet of *ten* lines. Where $\Delta m = 0$, the light vibrations occur in the direction of the field, so that the components underlined in the scheme are invisible in the longitudinal effect (*cf.* Fig. 74, where, as in Fig. 75, the lengths of the perpendiculars indicate the intensities of the lines, while, when directed upwards, they mark invisibility in the longitudinal effect).



FIG. 74.—Resolution of the *D*-lines of Na.

As a further example, we shall consider the anomalous resolution of the most important *line-triplet* of the elements in the second group, corresponding to the following transitions :—

$$^3S \rightarrow {}^3P_0, \quad ^3S \rightarrow {}^3P_1, \quad ^3S \rightarrow {}^3P_2$$

(*cf.* the previous Fig. 62). In the case of *mercury* the wavelengths of the components of this triplet are 4,047, 4,359, and 5,461 Å. It follows from Table XVI that the resolving factor for the $^3S$-term is 2, whilst (since $j = 1$) the possible values of $m$ are $(-1)$, 0, and $(+1)$. For the $^3P_0$-term the

[3] This line was originally called the *D*-line by Fraunhofer; the name has, of course, nothing to do with the *D*-terms.

resolving factor is indeterminate, and $m$ can only have the value 0 (since $j = 0$). Both for the $^3P_1$-term and the $^3P_2$-term the resolving factor, according to Table XVI, is 3/2 : the respective possible $m$-values are $(-1)$, 0, $(+1)$, and $(-2)$, $(-1)$, 0, $(+1)$, $(+2)$. Hence we have for the variations of the terms the scheme given in Table XVIII.

TABLE XVIII

*Term Variations for an Element in the Second Group*

| m= | −2 | −1 | 0 | +2 | +2 |
|---|---|---|---|---|---|
| $^3S$ | | −2 | 0 | +2 | |
| $^3P_0$ | | | 0 | | |
| $^3P_1$ | | $-\frac{3}{2}$ | 0 | $+\frac{3}{2}$ | |
| $^3P_2$ | −3 | $-\frac{3}{2}$ | 0 | $+\frac{3}{2}$ | +3 |

Let us now return to the combination rule, according to which $\Delta m = \pm 1$ or 0, and we shall assume that the intensity of a line vanishes when, on a transition, both $\Delta m$ and $\Delta j$ are equal to zero—an assumption which can be confirmed from considerations of atomic mechanics. We obtain then the following *line variations* for the triplet in question, the components which are not visible in the longitudinal effect being again underlined :—

$$(8)\begin{cases} ^3S \rightarrow {}^3P_0, & -2, & \underline{0}, & +2, \\ ^3S \rightarrow {}^3P_1, & -2, -\frac{3}{2}, & \underline{-\frac{1}{2}}, & \underline{+\frac{1}{2}}, & +\frac{3}{2}, +2, \\ ^3S \rightarrow {}^3P_2, & -2, -\frac{3}{2}, -1, & \underline{-\frac{1}{2}}, \underline{0}, \underline{+\frac{1}{2}}, & +1, +\frac{3}{2}, +2. \end{cases}$$

The three lines of the normal triplet therefore undergo respectively a three-fold, a six-fold, and an (equidistant) nine-fold resolution (*cf.* the scheme in Fig. 75). Photographs by *Back* of the resolution of this triplet for zinc are reproduced in illustrations *b* to *d* of Fig. 76 on Plate I. Illustration *a* shows the anomalous separation of the sodium doublet in the upper half of the photograph, whilst the lower half reproduces the doublet as seen in the absence of a field.



FIG. 75. Scheme of the anomalous Zeeman resolution of a line triplet.




FIG. 76.—Anomalous Zeeman effect: a, sodium doublet; b-d, zinc triplet; e, Paschen-Back effect in chromium triplet (after E. Back and A. Landé, "Zeeman-Effekt und Multiplettstruktur der Spektrallinien": Berlin; Springer, 1925).



FIG. 84.—Resolution of a silver ray in a magnetic field (after Gerlach and Stern).

[*To face p.* 130

On the gradual increase in strength of a "weak" magnetic field, the *anomalous* effect must, theoretically, *change by degrees into the normal Zeeman effect*. In point of fact, *Paschen* and *Back* in 1912 made the remarkable discovery that, when an originally "weak" field is gradually strengthened, the components of a resolved line begin to exert a peculiar influence on one another, just as if forces of attraction and repulsion were working between them. Preston's rule loses its validity, and the observed symmetry of the resolution disappears. This phenomenon is clearly to be seen in illustration *e* of Fig. 76 (Plate I.), which represents the so-called Paschen-Back effect for a line triplet of chromium: the upper and lower halves of the photograph show the components which vibrate respectively at right angles and parallel to the direction of the field.

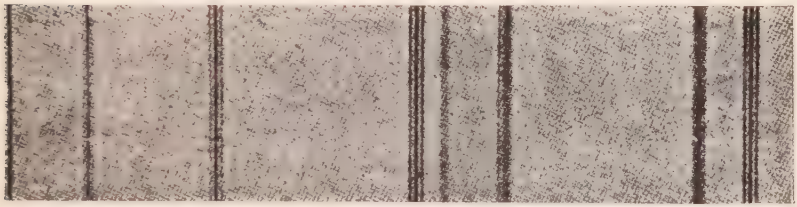


FIG. 77.—Normal Zeeman effect (after Zeeman).

As the strength of the field is increased, the Paschen-Back effect is finally transformed into the *normal* Zeeman effect, in which, as the photograph in Fig. 77 shows, the individual lines are resolved into equidistant triplets. The field-strength required for the production of the normal effect differs, of course, for different lines. For the sodium doublet it would amount to about 180,000 gauss—a strength which it has not yet been possible to generate for any length of time.[4] The first indication of the Paschen-Back effect, however, can be detected with the sodium doublet at about 30,000 gauss. An especially close investigation of the gradual transition from the anomalous to the normal effect has been possible with the *red lithium doublet* (6,708 Å.), the components of which differ by only 0·13 Å.

In a *strong magnetic field*, the term variation referred to

[4] Admittedly *Kapitza* has recently succeeded in producing fields of a few hundred thousand gauss, but only for a duration of time of about 1/100 second.

the normal separation as unit is described [according to § 102 eqn. (20)] by the formula

$$m^* = m_l + 2\,m_r. \tag{9}$$

In this expression, $m_l$ and $m_r$ are the components of the vectors **l** and **r**, and, as previously stated, $m_l$ can assume all integral values from $(-l)$ to $(-l)$, while $m_r$ may have any value at unit intervals between $(-r)$ and $(+r)$. The magnetic quantum number in a strong field, denoted by $m^*$, therefore differs from that in a weak field, for which we shall retain the simple symbol $m$. For, since $m$ may assume all values between $(-j)$ and $(+j)$, and $j$ corresponds to the resultant of the vectors represented by the quantum numbers $l$ and $r$, we have, in contradistinction to eqn. (9),

$$m = m_l + m_r. \tag{10}$$

In the case of alkali terms, $m_r$ can be only $(+\frac{1}{2})$ or $(-\frac{1}{2})$. For an alkali $S$-term $l = 0$, $r = \frac{1}{2}$; so that $m = (-\frac{1}{2})$ or $(+\frac{1}{2})$ and $m^* = (-1)$ or $(-1)$. For a $^2P$-term since $l = 1$, we may have $m_l = (-1)$, 0, or $(+1)$ and $m_r$ again $= (-\frac{1}{2})$ or $(+\frac{1}{2})$. Thus we have the following possibilities for the $^2P$-term :—

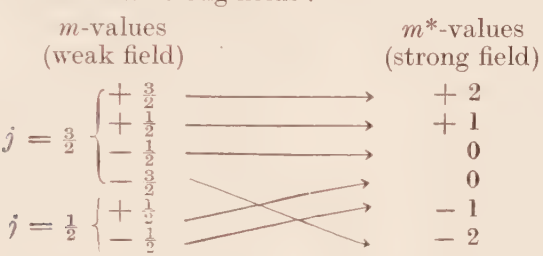
$$\left\{\begin{array}{lccccccc} m_l\,m_r, & -1,-\frac{1}{2} & 0,-\frac{1}{2} & +1,-\frac{1}{2} & -1,+\frac{1}{2} & 0,+\frac{1}{2} & +1,+\frac{1}{2} \\ m \text{ (weak field)}, & -\frac{3}{2} & -\frac{1}{2} & +\frac{1}{2} & -\frac{1}{2} & +\frac{1}{2} & +\frac{3}{2}\,; \\ m^* \text{ (strong field)}, & +2 & -1 & 0 & 0 & +1 & +2. \end{array}\right. \tag{11}$$

On writing the $m^*$-values under the appropriate $m$-values, we obtain

$$\begin{array}{lcccc} m = & -\frac{3}{2} & -\frac{1}{2} & +\frac{1}{2} & +\frac{3}{2}, \\ m^* = & -2 & -1, 0 & 0, +1 & +2. \end{array} \tag{12}$$

The question now arises as to the *correspondence* between the terms in a strong magnetic field and those in a weak field. We have seen that, in a weak field, the $^2P$-term undergoes a six-fold resolution; the lower component of the doublet $(j = \frac{1}{2})$ is associated with the $m$-values $(-\frac{1}{2})$ and $(+\frac{1}{2})$, and the higher component $(j = \frac{3}{2})$ with the $m$-values $(-\frac{3}{2})$, $(-\frac{1}{2})$, $(+\frac{1}{2})$, and $(+\frac{3}{2})$. Since it is only for the $j$-value $\frac{3}{2}$ that the values of $(\pm\frac{3}{2})$ and $(\pm\frac{3}{2})$ are possible for $m$, the terms $(m^* = \pm 2)$ and $(m^* = \pm 2)$ must undoubtedly be associated with the higher component $(j = \frac{3}{2})$ of the doublet. As regards the other four $m^*$-values (two of which coincide as zero values), we must make an assumption which may

readily be generalized. This assumption, which was propounded by *Sommerfeld* and subsequently more accurately established by *Pauli*,[5] expresses that, where two $m^*$-values are to be associated with one $m$-value, the lower $m^*$-value corresponds to the lower component of the doublet. In this way we obtain the following scheme for the association of terms in weak and strong fields :— [6]




Since in the normal Zeeman-effect the resolution of the *lines* (as opposed to resolution of terms) yields always only a *triplet*, one and the same line component must then in general be produced by *several different transitions*. The *distortion of the line separation in fields of medium strength* is explained by the fact that the critical value of the field-strength, at which the transformation of the anomalous into the normal effect begins, *differs* in general for the two terms which, by their combination, produce a certain line.[7]

## § 114. The Stark Effect.

The *resolution of spectral lines in an electric field* was discovered by *Stark* in 1913, and is therefore generally known as the *Stark effect* (Fig. 78).[1] The *theory* of the effect was developed from the *quantum principle* by *Schwarzschild* and *Epstein* in 1916. Between the results of theory and experi-

[5] W. Pauli, *Zeitschr. f. Phys.*, **20** (1923), p. 371.

[6] It should be noted that, in this arrangement and also in a more generalized form, terms with the same $m$-values do not intersect.

[7] For all details *vide* in particular E. Back and A. Landé, "*Zeeman-Effekt und Multiplettstruktur der Spektrallinien*" (Berlin ; Springer, 1925).

[1] The discovery of the Stark effect only succeeded after the skilful surmounting of great experimental difficulties. The observation of the resolution of the lines is, in fact, only possible in strong fields, which cannot be produced in a discharge tube, because it is too good a conductor of electricity. Stark therefore used canal ray particles to produce the spectral lines to be resolved, while he introduced a subsidiary electrode into the immediate neighbourhood of the cathode emitting the canal rays, thereby producing a very large potential difference over the short interval of a few millimetres. In a field of 30,000 volts/cm., the resolution of the line $H_\beta$ extends over 13 Å.

ment such an excellent agreement was found to exist that the Stark effect must be regarded as one of the most brilliant empirical confirmations of the quantum principle. The quantum theory of the Stark effect is, however, too difficult to be described in detail here, and we shall only state and discuss the result derived from the theory. According to this, the frequency ($\nu$) of a hydrogen line in an electric field of strength ($E$) is given by the expression

$$\nu = \nu_0 \pm z\, E\, \sigma, \tag{1}$$

where $\nu_0$ is the frequency of the line in the absence of an external field, and $\sigma$ is a universal constant of the form

$$\sigma = \frac{3\, h}{8\, \pi^2\, e\, m}, \tag{2}$$

$h$ denoting the elementary quantum of action, $e$ and $m$ the charge and mass of an electron. The "*ordinal number*" $z$,



Fig. 78.—Stark effect in the hydrogen lines $H_\delta$, $H_\gamma$, and $H_\beta$.

which is always an *integer*, depends on the principal quantum numbers $n$ and $s$ which characterize the initial and final states on the emission of the spectral lines; it is given by the relation

$$z = n\, n' \pm s\, s', \quad \begin{cases} n' = 0, 1, 2 \ldots (n-1), \\ s' = 0, 1, 2 \ldots (s-1). \end{cases} \tag{3}$$

The numbers $n'$ and $s'$ can thus assume all integral values between zero (and including it) and $n$ and $s$ (but excluding them) respectively.

The theory also leads to a rule governing the direction of vibration in the light corresponding to the spectral lines. If the expression

$$(n - s) + (n' - s') = k \tag{4}$$

is an *even* number, the *electric vibrations* are *parallel* to the direction of the electric field-strength. On the other hand, if $k$ is an *odd* number, the electric vibrations will occur *transversely* to the direction of the field-strength. As in the Zeeman effect, there are thus *two different classes of components* to be distinguished, which may be called the $p$- and the $t$-components from the initial letters of the words "parallel" and "transverse"; in the longitudinal effect the former are invisible.

The results of the theory may be discussed by reference to two simple *examples*, *viz.*, the *two violet hydrogen lines* $H_\gamma$ and $H_\delta$. As has already been mentioned, the line $H_\gamma$ is produced as third line of the Balmer series by a transition from a five-quantum to a two-quantum state. Hence, for the line $H_\gamma$, $n = 5$ and $s = 2$, so that $n'$ can assume the values 0, 1, 2, 3, 4, but $s'$ can only be 0 or 1. The ordinal number $z$ is therefore equal either to $5n'$ or to $(5n' \pm 2)$. In the first case the number $k$, as defined by eqn. (4), is equal to $(3 + n')$; in the second case, it is equal to $(2 + n')$. The ordinal numbers of the $p$-components, for which $k$ must be even, would thus be $5n'$ for odd values of $n'$, and $(5n' \pm 2)$ for even values. Hence the $p$-components of the line $H_\gamma$ would have the ordinal numbers

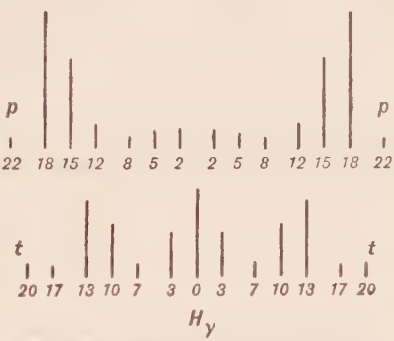



FIG. 79.—Stark effect in the line $H\gamma$.

$$2, 5, 8, 12, 15, 18, 22.$$

On the other hand, the ordinal numbers of the $t$-components, for which $k$ must be odd, would be $5n'$ for even values of $n'$, and $(5n' \pm 2)$ for odd values. Thus the $t$-components of the line $H_\gamma$ would have the ordinal numbers

$$0, 3, 7, 10, 13, 17, 20.$$

The complete *agreement* between theory and experiment is to be seen in Fig. 79, in which are reproduced the results of Stark's observations, the length of each line representing the intensity of the component in question, as estimated by Stark from the image on the photographic plate.

The fourth line in the Balmer-series, $H_\delta$, arises on a

transition from a six-quantum to a two-quantum state. For this line we have thus to put $n = 6$ and $s = 2$, so that $n'$ can have the values 0, 1, 2, 3, 4, 5, whereas $s'$ can again be only 0 or 1. The ordinal number $z$ is hence equal to either $6n'$ or $(6n' \pm 2)$. In the first case the number $k$ is equal to $(4 + n')$; in the second it is equal to $(3 + n')$. The ordinal numbers of the $p$-components, for which $k$ must be even, would therefore be $6n'$ for even values of $n'$, and $(6n' \pm 2)$ for odd values. Hence the $p$-components of the line $H_\delta$ would have the ordinal numbers

0, 4, 8, 12, 16, 20, 24, 28, 32.

On the other hand, the ordinal numbers of the $t$-components, for which $k$ must be odd, would be $6n'$ for odd values of $n'$, and $(6n' \pm 2)$ for even values. The $t$-components of the line $H_\delta$ would therefore have the ordinal numbers

2, 6, 10, 14, 18, 22, 26, 30.

As may be seen from Fig. 80, the agreement between theory

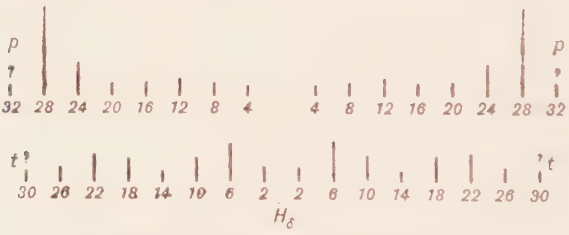



FIG. 80.—Stark effect in the line $H_\delta$.

and the previously known experimental results is excellent also in the case of the line $H_\delta$.[2]

For the simplest case of the Stark effect (considered above), in which the effect is "*linear*," *i.e.*, approximately proportional to the field-strength, quantum mechanics yields the same result as the older quantum theory. On the other hand, the newer theory gives a better agreement with experiment for the non-linear effect, such as is observed with the alkali spectra.[3]

[2] The occurrence of the components with the highest ordinal numbers is, it is true, doubtful in the experimental results, whilst the line of zero ordinal number is absent, possibly because its intensity is too small.

[3] *Cf.* especially the researches by Y. Ishida and his collaborators in *Scient. Papers Inst. Phys. Chem. Res. Tokyo*, 9 (1928).

## § 115. Intensities of Spectral Lines.

In 1924, *Ornstein* and his collaborators made the important discovery that the *intensities of the components of a line multiplet* stand in *simple rational numerical proportions* to one another. It has been found that a relation, called the *summation rule of Ornstein, Burger*, and *Dorgelo*,[1] is very accurately obeyed. According to this rule, the sum of the intensities of the lines, associated with a given term as *common initial* or *final level*, is proportional to the "*weight*" of this term. The weight of a term is determined by the number of $m$-values possible for it, *i.e.*, by the number $(2j - 1)$, where $j$ is the inner quantum number of the term.

The Ornstein-Burger-Dorgelo summation rule, it is true, suffices for the determination of line intensities only in line triplets, and thus generally for *doublet terms*. If we consider two doublet terms with the subordinate quantum numbers $(l - 1)$ and $l$, the corresponding $j$-values are $(l - \frac{3}{2})$, $(l + \frac{1}{2})$—twice—and $(l - \frac{1}{2})$—once. Since the $j$-values $\frac{3}{2}$ and $(-\frac{1}{2})$ cannot combine, we obtain a triplet (as was shown in § 106), and this becomes a doublet in the special case in which $l = 0$. Denoting the intensities of the three lines by $a$, $b$ and $c$, as shown in Fig. 81, we find from the summation rule, on the one hand, the ratio

$$a : (b + c) = (l + 2) : (l + 1) \tag{1}$$

and, on the other hand,

$$(a + c) : b = (l + 1) : l\,. \tag{2}$$

In the particular case of the *diffuse subordinate series* (transition from a $d$- to a $p$-term), we have to put $l = 1$, and we find

$$a : (b + c) = 3 : 2\,,$$
$$(a + c) : b = 2 : 1\,;$$

whence we obtain [2]

$$a : b : c = 9 : 5 : 1\,. \tag{3}$$

In point of fact, measurements carried out by Ornstein and his collaborators [3] have shown that this intensity ratio

[1] *Cf.* the papers by Dorgelo, Ornstein, and Burger in the *Zeitschr. f. Phys.*, **13** (1923), p. 206; **22** (1924), p. 170; **23** (1924), p. 258; **28** (1924), p. 241.

[2] A more general relation for the intensity ratio of the line triplet is $(2l^2 + l - 1) : (2l^2 - l - 1) : 1$, where $l$ is the larger of the two subordinate quantum numbers.

[3] The experimental methods, upon which are based the very numerous measurements of intensity performed in the Utrecht laboratory, were worked out by *Moll* and *Ornstein*.

holds very accurately for the triplets in the diffuse subordinate series of the alkali metals.

Since, as was previously mentioned (§ 106), the doublet separation for the $d$-term is smaller than for the $p$-term, the line denoted by $b$ in Fig. 81 must have the largest frequency, and thus the smallest wave-length, while the line $c$ will have the largest wave-length. In the triplets of the diffuse subordinate series, therefore, the *central line* is always the *strongest*, and the line of greatest wave-length the weakest.

In the case of the *principal series*, it follows from the summation rule that the two components of a line doublet (*cf.* Fig. 82) stand in the simple ratio of $(2j' + 1) : (2j'' + 1)$, where $j' = \frac{1}{2}$ and $j'' = \frac{3}{2}$. The intensity ratio

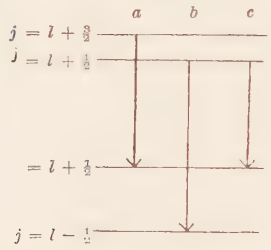



Fig. 81.

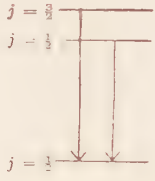



Fig. 82.

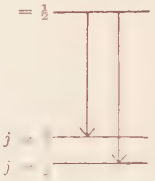



Fig. 83.

is therefore 1 : 2, the *component of shorter wave-length* being *twice as strong* as the other. The converse holds for the *sharp subordinate series* (Fig. 83), where the ratio is again simply 1 : 2, but the *component of greater wave-length* is the stronger.

When we have to deal with more than three components, the summation rule does not suffice for the determination of the intensities, and we have to take into consideration other matters which cannot be discussed here.[4] *Sommerfeld* had previously discovered two quite general rules which have proved most useful in the analysis of spectral terms. The components, for which the inner quantum number changes in the *same sense* as the subordinate quantum number, are always *stronger* than those lines for which the changes occur in opposite senses. Furthermore, the intensities of the components associated with the larger $j$-value are always stronger than those corresponding to a smaller $j$-value.

[4] *Cf.* H. Hönl, *Ann. der Phys.*, **79** (1926), p. 273.

For the rest, the summation rule as given above holds only *approximately*, *viz.*, only on the assumption that the intensity can be regarded as independent of the frequency values. This assumption is only justified for a small multiplet separation, and in other cases complicated corrections are necessary. With the application of these, the summation rule is admittedly obeyed in the majority of experimental observations.[5] Comparatively simple intensity ratios may be expected for the *components of the Zeeman* and the *Stark effect*. Both these problems were treated approximately prior to the foundation of the new quantum mechanics, but the latter first led to their exact solution.[6]

## § 116. The Magneton.

Substantial support is afforded to the fundamental hypotheses of modern atomic theory by their confirmation, quite independently of all spectroscopic experimental results, by observations of a purely electrical, a purely chemical, and a purely magnetic character. Attention has been drawn to electrical observations in connection with the phenomenon of electron collisions, and we have frequently pointed out support of a chemical character when discussing the periodic system (*e.g.*, the parallelism between the chemical bivalency and the spectroscopic ambiguity of copper). Purely *magnetic observations*, which represent an extremely important confirmation of the fundamental principles of atomic theory, were obtained in 1922 by *Gerlach* and *Stern* with the so-called *atomic rays*.

These two investigators subjected to the action of a *non-homogeneous magnetic field* the rays of atoms produced on the *vaporization of metals* in a suitable apparatus,[1] thereby causing a *deflection of the rays*.[2] According to classical

[5] *Cf.* the summary by L. S. Ornstein in the *Physikal. Zeitschr.*, **28** (1927), p. 688.

[6] *Vide* M. Born's "*Problems of Atomic Dynamics*" (Cambridge, U.S.A.; 1926), Series I., lect. 17, and E. Schrodinger's "*Collected Papers on Wave Mechanics*" (London; Blackie, 1928).

[1] For further details of this experiment and of other problems here discussed, *vide* W. Gerlach's report on atomic rays in the "*Ergebnisse der exakten Naturwissenschaften*" (Berlin, Verlag Springer), **3** (1924), p. 182. The discoverer of atomic rays was Anthony (1892).

[2] A homogeneous magnetic field produces no deflection because, in it, the positive and negative poles of a magnet are attracted with equal forces in opposite directions. In experiments on atomic rays, the gradient of the magnetic field-strength amounts to some hundred thousand gauss per centimetre.

theory, this deflection must be *proportional to the component of the magnetic moment of the individual atom in the direction of the gradient* of the field-strength.

Now it follows from the quantum theory, as has been explained in a previous section (§ 102), that an atom cannot assume any arbitrary position in a magnetic field, but that it can orientate itself only in definite "*quantized directions*." As we have already seen, the number of possible orientations is equal to $(2j + 1)$, where $j$ is the *inner quantum number* characteristic of the normal state of the atom. Gerlach and Stern's first experiments were concerned with *rays of silver*. Since the fundamental term of silver is a doublet $S$-term, *i.e.*, a term with a $j$-value of 1/2, *two different orientations* are apparently possible for the silver atom. In point of fact, in Gerlach and Stern's experiment the silver beam, the cross-section of which measured a few hundredths of a millimetre, was *split into two discrete rays* in the non-homogeneous magnetic field, so that a direct *experimental proof of quantized orientations* was thus obtained (*vide* Fig. 84 on Plate I). From the magnitude of the deflection it was possible to *calculate the component of the magnetic moment* associated with a *single atom of silver*, and the value found was $9 \times 10^{-21}$ absolute units.

Quite generally, we can derive a simple *relation between the mechanical angular momentum and the magnetic moment* to be ascribed to an arbitrary revolving body when it is electrically charged. If $Q$ be the charge and $\omega$ the revolution frequency, the current is given by the product $Q\omega$ in electrostatic units, and by the expression $Q\omega/c$ in electromagnetic units. Hence, by § 60 eqn. (1), we find for the magnetic moment

$$M = \frac{Q\,\omega\,F}{c}, \tag{1}$$

where F denotes the area of the orbit, or circuit area. On the other hand, in the case of a circular orbit of radius $a$, the mechanical angular momentum is given by the relation

$$U = m\,a\,v = m\,a\,.\,2\,\pi\,a\,\omega = 2\,m\,\omega\,F. \tag{2}$$

Hence

$$M = U\,\frac{Q}{2\,m\,c}\,. \tag{3}$$

If, now, we define as a *magneton* ($\mu_0$) the value of the magnetic moment for a mechanical angular momentum of

magnitude $h/2\pi$, and for the charge and mass of an *electron*, we find

$$\mu_0 = \frac{e\,h}{4\,\pi\,m\,c}, \tag{4}$$

or

$$\mu_0 = 9{\cdot}21 \times 10^{-21} \text{ abs. units.} \tag{5}$$

It follows therefore from Gerlach and Stern's researches that an *atom of silver* takes up an orientation in a magnetic field such that the component of its magnetic moment amounts to *exactly one magneton*.

The quantum theory leads us to expect that, with an arbitrary element, the deviations in the deflected atomic ray will correspond to the products $g\ m$, where $g$ denotes *Landé's resolving factor*, and $m$ the possible values of the magnetic quantum number. With *silver*, for which the fundamental term is a $^2S_{1/2}$ -term, we have (by Table XVI), $g = 2$, and, since $j = \frac{1}{2}$, the possible values of the magnetic quantum number are $(+\frac{1}{2})$ and $(-\frac{1}{2})$. An atomic ray experiment with silver must therefore actually yield separations of $(+1)$ and $(-1)$, referred to the magneton as unit. Identical results were naturally obtained in experiments with copper and gold rays, since the fundamental term for both these elements is of the same character as for silver.

Inasmuch as terms with a zero inner quantum number are not resolved, we should expect *no deflection* at all with atomic rays of those elements which have a fundamental term with $j = 0$. This is the case for elements in the second group ($^1S_0$ -term) and for those in the fourth group ($^3P_0$-term). In point of fact, experiments by Gerlach, Stern, and their collaborators showed that there is no deflection with the second group elements zinc, cadmium, and mercury, and also with the fourth group elements tin and lead. For *thallium*, an element in the third group, the fundamental term is a $^2P_{1/2}$ -term, so that the resolving factor is equal to $\frac{2}{3}$. The possible $mg$-values are therefore $(+\frac{1}{3})$ and $(-\frac{1}{3})$, and equal and opposite values for the deflection were actually obtained by experiment, corresponding to one-third of a magneton.

Since the inner quantum number $j$ determines the total angular momentum of the atom, atoms of elements which have a fundamental term with $j = 0$ can possess no intrinsic magnetic moment: such elements must therefore be *diamagnetic* in character. *Paramagnetism* is only possible

when, in the normal state, $j$ differs from zero. In a magnetic field, the atoms of a paramagnetic substance naturally endeavour to take up *parallel* orientations, but their efforts are opposed by the incessant *molecular collisions* consequent on the internal heat motion, and the attainment of a "magnetic saturation" is thus defeated. As a result, a kind of statistical equilibrium is produced, and this is manifested in a *dependence of paramagnetism on temperature*. Experiment proves that, for every paramagnetic substance, the *magnetic susceptibility* (*viz.*, the magnetic moment produced by a field-strength of 1 gauss) is *inversely proportional to the absolute temperature*. The product of the susceptibility, referred to a gram-atom, and the absolute temperature thus represents for any given element a characteristic constant which is called its *Curie* constant.[3] On the assumption (which at the time seemed self-evident) that *all orientations* were *equally possible* for an atom in a magnetic field, *Langevin* in 1905 derived for the Curie constant the relation

$$C = \frac{1}{3}\frac{\mu^2 L^2}{R}. \tag{6}$$

In this formula $L$ denotes the Loschmidt number, $R$ the absolute gas constant ($8{\cdot}313 \times 10^7$ ergs deg.), and $\mu$ the magnetic moment of a single atom (or a single molecule, if the susceptibility be referred to a gram-molecule).

When, in accordance with the newer theory, *direction* is to be *quantized*, the numerical factor 1/3 has to be replaced by the mean value of $m^2/j^2$, where $m$ is the magnetic quantum number, which can assume all values at unit intervals from $(-j)$ to $(+j)$. Thus we find

$$\overline{\left(\frac{m}{j}\right)^2} = \frac{1}{j^2}\,\frac{1}{2j+1}\sum_{m=-j}^{m=+j} m^2. \tag{7}$$

According to a well-known algebraical relation,

$$1^2 + 2^2 + 3^2 + \ldots + x^2 = \frac{x(x+1)(2x+1)}{2.3},$$

and it may easily be shown that the same result is obtained for the sum $(\frac{1}{2})^2 + (\frac{3}{2})^2 + \ldots + x^2$, where $x$ is an odd multiple of 1/2. Hence

$$\sum_{-j}^{+j} m^2 = \frac{j\,(j+1)\,(2j+1)}{3}, \tag{8}$$

[3] This constant is so called because it was Pierre Curie who discovered the inverse proportionality between susceptibility and temperature.

and, on inserting this value in eqn. (7), we obtain

$$\overline{\left(\frac{m}{j}\right)^2} = \frac{j+1}{3j}. \tag{9}$$

Finally, in the sense of the newer theory, we have to put

$$\mu = g\, j\, \mu_0 \tag{10}$$

in eqn. (6), where $\mu_0$ again denotes the magneton. On replacing the factor 1/3 in eqn. (6) by the factor given in eqn. (9), we find therefore

$$C = \frac{1}{3}\frac{\mu_0^2 L^2}{R} g^2 j\,(j+1). \tag{11}$$

If we replace the universal constants $\mu_0$, $L$, and $R$ by their respective values, we obtain in this way the important formula

$$g\sqrt{j\,(j+1)} = 2{\cdot}83\sqrt{C}\,. \tag{12}$$

It has been proved that this relation is obeyed fairly exactly by paramagnetic elements.[4]

It is evident from eqn. (10) that the definite magnetic moments of a single atom are indeed rational, but that generally they need *in no way* be *integral multiples* of a magneton. This fact removes an obstacle which, in earlier years, had presented a great difficulty to theoretical physicists. Even in 1911, before the establishment of Bohr's atomic theory, *Pierre Weiss* thought that he had proved that, in a state of magnetic saturation (thus at sufficiently low temperatures), the magnetic moments of the individual molecules must be integral multiples of a certain quantity. Weiss called this quantity a magneton, but the value which he derived for it from empirical data was only *one-fifth* of the value subsequently calculated by *Bohr* for the quantity $\mu_0$, on a theoretical basis.[5] This apparent contradic-

[4] In this connection special attention may be drawn to *Hund's* attempt to derive the Curie constant from eqn. (12), with the help of some reasonable spectroscopic assumptions, for the *three-fold ions of the rare earths* (which presumably contain only one to fourteen *f*-electrons outside the xenon core). Except in two cases, No. 62 (samarium) and No. 63 (europium), Hund found an excellent agreement with the available experimental values. *Cf.* Hund, *l.c.*

[5] Weiss believed that, in a state of saturation, the magnetic moments referred to the gram-molecule must be integral multiples of 1123 abs. units. Division by the Loschmidt number yielded the value $1{\cdot}85 \times 10^{-21}$; this quantity was formerly known as Weiss's magneton in contradistinction to Bohr's magneton, which is 4·97 times as large. In this connection it may be mentioned that, in *ferromagnetic* bodies, which exhibit a permanent magnetism, theory assumes the presence of forces acting between the molecules. These

tion was, however, solved by the theoretical considerations quoted above, which are principally due to *Pauli*.

The concept of a *magneto-mechanical parallelism* underlying the modern theory of magnetism received a direct confirmation from two experiments, one of which was carried out by *Einstein* and *de Haas*, and the other by *Barnett*. Einstein and de Haas, in 1915, succeeded in causing a soft iron rod to *rotate by rapid reversals of magnetism*, and the converse effect was produced in 1917 by Barnett, who succeeded in *magnetizing* a rod of *soft iron by rapid rotation*. A perfect quantitative agreement was found for both effects between observed and calculated values, apart from the presence of a factor 1/2, and an explanation was subsequently provided for this in the spin of the electrons.

### § 117. The Excitation of Atoms.

The state of *excitation* represents for atoms an abnormal, extremely transient condition, into which they can be brought by the absorption of light quanta. A far more frequent way of exciting atoms than by absorption, however, is by supplying the requisite energy from the *heat store* of the body concerned. According to the fundamental concept of the kinetic theory of heat, incessant *collisions* occur in a gas between the rapidly moving molecules. By virtue of these collisions, *energy of translation* can of course be *changed into intra-atomic energy*, but when the atom struck was previously in the normal state, this can only happen if the relative kinetic energy is at least as large as the excitation energy. (In the monatomic vapours of the metals in the first two groups, which have been chiefly investigated, as well as in the rare gases, the concepts of atom and molecule are synonymous.) Now, when atoms thus excited by collisions return to the normal state with the emission of light-quanta, the well-known phenomenon of *thermo-luminescence* is observed. Moreover, the atoms of a gas can be excited, not only by sufficient heating, but also by collisions with rapidly moving ions *in electrical fields*, as is shown by the *luminescence of rarefied gases in discharge tubes*.

The *mean energy* possessed by a monatomic *gas molecule* at an absolute temperature $T$, in virtue of its *translation*, is, as follows from the kinetic theory of gases (and as we shall

forces become ineffective, however, above a certain temperature, known as the *Curie point*, at which a discontinuity in the specific heat is also to be observed. The Curie point for iron, for example, lies at 750°, and for nickel at 375°. *Vide* also § 135.

show in detail in a subsequent section, *viz.*, § 165), equal to $3/2\ k\ T$, where $k$, the so-called Boltzmann constant, is equal to $1{\cdot}372 \times 10^{-16}$ ergs/deg.[1] At room temperature, for which $T$ is about 300, the mean value of the translational energy of a monatomic molecule is hence about $6 \times 10^{-14}$ ergs. Furthermore, according to a fundamental relation of molecular statistics, the *fraction* ($\alpha$) of molecules of which the translational energy is greater than $x$ times the mean value is, quite generally, equal to $e^{-x}$.

Thus, in the case of cæsium, which has the smallest excitation energy of all the alkali metals, *viz.*, $2 \times 10^{-12}$ ergs at room temperature, the fraction ($\alpha$) of the atoms of which the translational energy exceeds the excitational energy is about $10^{-13}$ (for $x$ is approximately equal to 30). At a temperature of 1,200° C. (= 1,500° abs.) $x$ becomes roughly equal to 6, and hence $\alpha$ to about 1/400. Whereas, therefore, at room temperature the fraction of atoms possessing a kinetic energy capable of exciting another atom by collision is entirely insignificant,[2] at 1,200° C. this fraction is large enough to produce a noticeable luminescence in the vapour.

Theory teaches us to expect that, when any vapour commences to glow with rising temperature, the first spectral line to be observed will be that corresponding to the transition from the lowest excited state to the normal state. The longer the wave-length of this line, the smaller will be the excitation energy, and the lower the temperature at which emission of the line will begin. As this temperature is exceeded, we must further expect that the transfer of translational energy will be sufficient to produce higher and higher excitations, so that new spectral lines will be continually appearing. In point of fact, all these theoretical conclusions are confirmed by experiment.

The *ionization energy* proves in general to be two or three times as large as the excitation energy. Hence, if we now refer the fraction $\alpha$ to the process of ionization instead of to that of excitation, as previously, we should expect to find the same value of $\alpha$ for an absolute temperature two or three times as high. Actually, however,

[1] In the case of molecules of diatomic gases, the factor 3/2 is replaced by the factor 5/2. Boltzmann's constant is the quotient of the gas constant and the Loschmidt number (*cf.* § 165).

[2] Actually, it is not the mean translational energy, but the mean value of the relative energy on collision, that comes in question. It follows from the kinetic theory of gases, however, that the two quantities only differ by an inconsiderable amount.

ionization is to be observed for much smaller values of $\alpha$ than is excitation. For, on introducing the salts of a metal into a flame, the smallest ionization can at once be observed by a noticeable *electrical conductivity of the flame*, for the detection of which we possess in the galvanometer an extraordinarily sensitive instrument. As theory leads us to expect, the flame conductivity at a given temperature is greater the smaller the optical fundamental term of the metal concerned.

Ionized atoms can, of course, be excited in an entirely similar way to neutral atoms, by the transition of one of the remaining valency electrons to a higher-quantum state. The easiest spark spectra to produce are those of the second group elements, a temperature of 2,000° C. being sufficient for this purpose, for their simple ions still contain one loosely bound electron. Arc and spark spectra of these elements are thus often to be observed together. On the other hand, it is extremely difficult to generate the spark spectra of the alkali metals, since their atoms in the neutral state only contain a single valency electron.

These conclusions from atomic theory provide an explanation of several apparently paradoxical results in *solar spectroscopy*, as was first pointed out in 1920 by the Indian physicist, *Saha*.[3] It had previously been quite inexplicable why sodium should appear very strongly in the solar spectrum, whereas potassium appears only weakly, and rubidium and cæsium are entirely absent. According to Saha's striking explanation, the alkali metals are present in the sun, just as in the earth. *To some extent*, however, they *remain unnoticeable*, because, in consequence of the high temperature, they are almost invariably *ionized*, and their spark spectra, the lines of which lie in the ultra-violet, escape spectroscopic observation. That the arc lines of Na occur strongly in the solar spectrum, those of K only weakly, and those of Rb and Cs not at all, follows from the fact that Rb and Cs have the smallest ionization energies, K having a somewhat higher, and Na a considerably larger value.[4] In contrast to the alkali metals, the alkaline earths are represented in the visible portion of the solar spectrum by the lines of their ionized atoms as well; for instance, the strong

[3] M. N. Saha, *Phil. Mag.*, 40 (1920), pp. 472 and 809, and 41 (1921), p. 267. *Cf.* also Kohlschütter's report in *Naturwissenschaften*, 9 (1921), p. 863.

[4] *Cf.* Table VII in § 106.—Besides the differences in ionization energy, the abundance or rarity of the elements, it is true, also play a part.

Fraunhofer lines $H$ and $K$ both belong to the spark spectrum of calcium.[5]

Moreover, Saha found a simple explanation for the remarkable fact that the heavier element calcium is to be observed in the spectrum of the sun's atmosphere at much greater heights than the much lighter element hydrogen. For the calcium lines lying in the visible region of wavelengths require a smaller excitation energy than the hydrogen lines lying in the same region ; the hydrogen lines which are *readily* excited belong to the ultra-violet.[6] In the higher regions of the sun's atmosphere, where the temperature is less than in the lower regions, the energy transferred by molecular collisions therefore suffices to excite the visible calcium lines, but not the visible hydrogen lines.

The energy required to tear from an atom one of its *inner planetary electrons* is, as follows from a consideration of the *X-ray terms*, from some hundred to about 20,000 times as much for the highest elements as the ionization energy of, say, sodium. In the case of *uranium*, for example, we find from its X-ray spectrum that the ratios of its $K$-term and of its numerically largest $L$- and $M$-terms to the Rydberg term have the values 8,477, 1,603 and 409, whereas, for *sodium*, the ratio of the fundamental term (which is a measure of the ionization energy) to the Rydberg term amounts to 0·378. The ionization energy of the hydrogen atom is equal simply to the product of the Rydberg term, the velocity of light, and the elementary quantum of action, so that its value is $2{\cdot}15 \times 10^{-11}$ ergs. Equating this energy to $\frac{3}{2} kT$, we can calculate the *equivalent temperature* to be about 104,000°. For sodium, the equivalent temperature (obtained on multiplying by 0·378) would be about 39,000°. The three considered terms of uranium would correspond to equivalent temperatures of some 880, 170, and 40 million degrees. For intermediate elements, we should have to assume, in accordance with Moseley's law, temperatures about a quarter of that for uranium.

Now, on the basis of various considerations, astronomers have agreed in ascribing to the *temperature in the interior of a fixed star* a value of some *forty million degrees*.[7] According

[5] The earlier view that calcium is the most abundant element in the sun's atmosphere has, of course, been proved erroneous by Saha's researches.

[6] The hydrogen line corresponding to the excitation energy, in the narrower sense of the word, is the first line in the Lyman series (1,216 Å).

[7] *Cf.* A. S. Eddington's "*The Internal Constitution of the Stars*" (Cambridge : University Press, 1926), or his excellent popular work, "*Stars and Atoms*" (Oxford : Clarendon Press, 1928).

to what we have just said, it follows necessarily that, *in the interiors of stars*, the *atoms* must be *largely deprived of their planetary electrons.*

This deduction provides a very simple explanation for an apparently highly paradoxical conclusion to which astronomers had been led by a study of the *companion of Sirius.* They found that the surface of this star glows more strongly than that of the sun, but that the total radiation emitted amounts only to the 1/360th part of the solar radiation. Consequently, the surface of the companion of Sirius must be at least 360 times as small as the surface of the sun. On the other hand, however, astronomers could not possibly doubt the certain conclusion of celestial mechanics that the mass of the companion of Sirius is about four-fifths of that of the sun. If both these deductions are correct, the *density* of the companion of Sirius, as *Eddington* pointed out, must be some 60,000 times that of water, and thus more than 2,000 times the density of platinum. This very remarkable conclusion was completely confirmed when (as has already been mentioned in § 91) *Adams* actually succeeded in demonstrating for the companion of Sirius, as an "Einstein effect," the extraordinarily *large line displacement* corresponding to the assumed values of the radius and the mass. As Eddington showed, the riddle of such an *enormous stellar density* is solved by the recognition that, in the interior of stars, the atoms are almost entirely deprived of planetary electrons, so that the atoms thus truncated occupy a far *smaller volume* than the complete terrestrial atoms.

Just as the excitation of an atom may be produced by irradiation or by the addition of mechanical energy, so the *return of an excited atom to the normal state* may or may not be accompanied by the emission of radiation. When, in particular, excitation is due to irradiation, and radiation is emitted on the return, the well-known phenomenon of *fluorescence* arises. In the special case in which the wave-lengths of the secondary and the primary radiation are identical, we speak of a *resonance radiation.* According to the discussion in a previous section (§ 106), we may expect to meet with such radiation when light of the wave-length of a *resonance line* of an element passes through the vapour of the element at a temperature below that at which any of its atoms are excited.

The phenomenon of resonance radiation was discovered and investigated, long before the atomic theory was estab-

lished, principally by *E. Wiedemann* and *Wood* with *sodium vapour*, and by *Wood* in the case of *mercury vapour*. At low pressures the total energy of the re-emitted light does, in fact, agree with the energy absorbed from the primary radiation. It is only when the pressure is increased that a part of the absorbed primary radiation is changed into molecular energy of translation, with a resultant heating of the fluorescent vapour.

When a gas is irradiated with light of a frequency *greater than the resonance frequency* and corresponding to the difference in energy between the normal state and a *higher*, excited state, the *return* of the atom from this more highly excited state may take place *in steps*. In the fluorescent radiation, therefore, lines of greater wave-length than that of the exciting light may occur. Under no circumstances, however, can an atom which, before excitation, is in the normal state, emit fluorescent radiation of shorter wave-length than the primary radiation, and in this lies the explanation of *Stokes's rule of fluorescence*, which was mentioned in an earlier section (§ 91). Only when the irradiated atom has been previously excited can it happen, by way of exception (*e.g.*, on a return to the normal state), that the wave-length of the fluorescent radiation may be shorter than that of the exciting radiation.

Of the excited atoms in a fluorescent gas, only a certain number in general will return to the normal state with the emission of radiation, the rest returning without radiating consequent on their *collisions* with other atoms. In such cases the excess atomic energy must be transformed into translational energy, with a resultant increase in the heat energy present. The ratio between the two numbers evidently depends on the ratio between the *mean time between two collisions* (to be calculated from the theory of gases) and the *average duration of the excited state*. As long as the former interval is considerably greater than the latter, the total energy of the fluorescent radiation must be approximately as large as the total energy absorbed from the primary radiation by the fluorescing gas. If, however, the mean interval between two collisions be *shortened*, as by *increasing the pressure* by the *admixture of another gas*, the *fluorescence* should be gradually *extinguished* as the mean interval between two collisions becomes of the same order of magnitude as the average duration of the excited state.

The extinction of fluorescence was actually observed by *Wood*, when air was mixed with the fluorescent mercury

vapour.[8] Both his observations and those of later investigators [9] lead to an approximate value of $10^{-8}$ sec. for the *mean duration of the excited state* in mercury. One-tenth as large (as we may remark in passing) is the time found by *Wien* for the *decay of the luminescence* of excited *canal ray particles*.[10]

Moreover, it has been found possible to demonstrate experimentally a *transference of excitation energy* from the *atoms of one substance* to those of *another*, the transference being obviously effected by collisions. This phenomenon, in which light is absorbed by atoms of one kind and re-emitted by atoms of another kind, is known as *sensitized fluorescence*. It was first observed, in 1922, by *Franck* and *Cario*,[11] in a *mixture of mercury vapour and thallium vapour*. This mixture was irradiated at suitable pressure and low temperature with ultra-violet mercury light of the resonance wave-length of 2,537 Å. These rays are naturally absorbed by the mercury atoms, but not by the atoms of thallium, the spectrum of which exhibits in the neighbourhood of 2,500 Å. no lines having the normal state as initial state for absorption. None the less, the spectrum of the fluorescent radiation contained all the lines of thallium, the excitation energy of which is smaller than a light-quantum of 2,537 Å. wave-length.

The time during which an atom *persists* in an excited state naturally depends on the *probability of a transition* from this to another state. Comparatively *long durations in the excited state* may thus be expected when an atom is displaced, perhaps by a collision, into an abnormal state, from which a *return* to the normal would be *precluded* by the spectroscopic combination rules (although these do admit of exceptions). A state of this kind is termed *metastable*. Generally, it will be destroyed by collisions with other atoms long before the "forbidden" line can be emitted.

Metastable states have actually been detected with the rare gases, with mercury, and also with some other elements, and their durations have been measured. For example, *Meissner* [12] discovered a metastable state for *neon* with

[8] Wood, *Phil. Mag.*, **21** (1911), p. 309.

[9] *Cf.* G. Cario, *Zeitschr. f. Phys.*, **10** (1922), p. 185.

[10] W. Wien, *Ann. der Phys.*, **60** (1919), p. 597, and **66** (1921), p. 229.

[11] *Cf.* the summary by J. Franck in the second volume of "*Ergebnisse der exakten Naturwissenschaften*" (Berlin: Springer, 1923).

[12] K. W. Meissner, *Ann. der Phys.*, **76** (1925), p. 124. Meissner's experiments consisted essentially in passing an alternating current through two discharge tubes filled with the gas under investigation, one tube serving for the emission, and the other for the absorption, of the lines.

mean durations of 0·0036 and 0·0042 sec. until it was terminated by the emission of two forbidden lines: for *argon* one persisting for about 1/75 sec. was found by *Dorgelo*.[13] In these instances, therefore, the durations of the metastable states are about a hundred thousand times those of ordinarily excited states.

The *metastability of helium* is particularly noteworthy. By means of *electron collisions*, *Franck* has succeeded in transforming helium atoms from the normal para-state into the lowest ortho-state (*cf.* § 107), for which process a so-called *transformation potential* of 19·8 volts is necessary. The actual metastability of the two-quantum ortho-state thus produced follows from the fact that this state is associated with a characteristic *resonance line*. This line corresponds to the transition from an ortho-$P$-term to the normal ortho-$S$-term, and its wave-length is 10,830 Å., or approximately 1 $\mu$. If helium in which a sufficient number of metastable atoms has been produced by electron collisions be irradiated with infra-red light of about 1 $\mu$ wave-length, perfect resonance ensues.

The metastable state is, of course, associated with a *smaller ionization energy* than the normal state. The difference between the ionization potentials of the normal and the metastable atom must, indeed, be equal to the transformation potential. Since the ionization potential of normal helium is 24·5 volts, and the transformation potential 19·8 volts, the ionization potential of metastable helium must amount to 4·7 volts. The metastable helium atom must therefore lose an electron more easily than lithium and sodium,[14] and hence must be still *more electropositive* than these alkali metals, and, in contrast to the rare gases, it should exhibit considerable *chemical activity*. In point of fact, *Manley* [15] has succeeded in demonstrating the formation of *compounds between helium and mercury* (probably $HgHe_2$ and $HgHe_{10}$); while the occurrence of a band spectrum of helium at sufficiently high densities may be referred to the formation of molecules by metastable atoms.

Lines debarred by spectroscopic principles cannot as a rule be observed, because, even with the greatest producible attenuation, an atom cannot remain undisturbed for more than a fraction of a second. Conditions are, how-

[13] *Cf.* H. B. Dorgelo and T. P. K. Washington, *Proc. Roy. Acad. Amsterdam*, **30** (1927), p. 33. *Vide* also the report on the duration of luminescence in atoms by Kerschbaum, *Ann. der Phys.*, **83** (1927), p. 287.

[14] *Cf.* Table VII. in § 106.

[15] *Cf.* J. J. Manley, *Phil. Mag.*, (7) **4** (1927), p. 699.

ever, very different in the exceedingly attenuated *cosmic nebulæ*, where, as astrophysical considerations show, an atom may persist for a year quite undisturbed, Thus, in the atoms of cosmic nebulæ, there is an evident possibility of the accomplishment of forbidden transitions which, in terrestrial conditions, would be precluded by the incessant disturbances. In point of fact, astrophysicists have long been struck by the appearance in nebular spectra of lines which cannot otherwise be observed ; these lines were formerly ascribed to an element not known on the earth—the so-called *nebulium*.[16] In 1927 *Bowen*[17] succeeded in identifying the mysterious nebulium with *air*. As he showed, the most important " nebulium " lines correspond to forbidden transitions from metastable states of the *doubly ionized oxygen* and of the *ionized nitrogen atom*. On the plausible supposition that the total mass of oxygen in a nebula approximates to the mass of the sun, Eddington[18] proved that one such forbidden transition in a century, on the part of each oxygen atom, would suffice to generate a total nebular radiation exceeding the total solar radiation a hundred-fold. Furthermore, the mysterious *green auroral line* which, because it could not be identified with any known line, formerly gave rise to the most diverse speculations, has been recognized by *McLennan*[19] as a forbidden line of *oxygen*.

[16] This hypothesis was apparently justified by the circumstance that helium was first detected in the sun, and only subsequently discovered on the earth : it became untenable on the establishment of the natural sequence of the elements.

[17] I. S. Bowen, *Nature*, **120** (1927), p. 473.

[18] *Vide* Eddington's " *Stars and Atoms* " (Oxford : Clarendon Press).

[19] J. C. McLennan, *Proc. Roy. Soc.*, *A*, **108** (1925), p. 501, and **120** (1928), p. 327.

## Chapter XIV

# X-RAYS

### § 118. Space Lattices

Soon after the discovery in 1895 of *X-rays*, physicists came to the conclusion that the new rays must consist of *electromagnetic waves of extremely short wave-lengths*. Various considerations led to the correct assumption of an order of magnitude of about $10^{-9}$ cm. for the wave-length,[1] and this also explained the failure of all early experimental efforts to demonstrate *interference* or *diffraction* phenomena with X-rays, by means of any ordinary optical method. The usual diffraction gratings could not be applied, for, as follows from the theory of the grating (and as we shall show in detail), the interval between two lines of the grating must not be much larger than the wave-length of the light investigated. Hence, researches on X-rays would require gratings upon which there were millions of lines to a millimetre. The technical impossibility of preparing such fine gratings seemed to exclude almost every prospect of physicists ever being able to determine the exact wave-lengths of X-rays, until, in 1912, *Laue* conceived the ingenious idea of making use simply of *crystals*, instead of artificial gratings, for diffracting the rays. Laue's idea was based on a theory developed by the mineralogist *Bravais* about 1850, according to which the regular form of crystals must be due to a *regular arrangement of the atoms in the crystal*. It followed also from the known density of crystals, and from the known value of the Loschmidt number, that the mutual distances between the atoms in crystals must be of the order of $10^{-8}$ cm. Provided that the individual atoms behave as diffraction centres, it was therefore to be expected that *crystals* would actually *behave as gratings when used with X-rays*.

In order to understand the behaviour of crystals with

[1] One such consideration was the supposition that an electron, which contributes to the production of X-rays, is slowed down along the portion of its path inside an atom.

respect to X-rays, we shall first consider, as the simplest of all gratings or lattices, a *line-lattice* in which the lattice points are arranged at equal intervals along a straight line. The constant interval between two successive lattice points is called the *lattice constant.*

We suppose now a *wave* to be incident on the lattice, the direction of the ray making with the direction of the lattice an angle the cosine of which is $\alpha_0$.[2] The individual lattice points may be regarded as diffraction centres, each of them constituting a centre of excitation of new waves. These waves travel in all directions, but it will be immediately evident that they will be *diffracted with especial intensity in those directions*, for which the *difference in path* of the rays

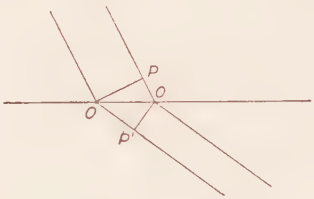



Fig. 85.

proceeding from two successive lattice points amounts to an *integral multiple of a wave-length.* In fact, more accurate theoretical considerations (which we cannot discuss here) show that, in consequence of mutual *interference* between the diffracted rays, an *appreciable* diffraction of the light occurs only in the directions just given.

We shall now denote by $\alpha$ the cosine of the angle included between the direction of the lattice and one of the rays thus diffracted. Then, if $a$ be the lattice constant, the difference in path of the rays diffracted in this direction by two neighbouring lattice points is equal to $a\ (\alpha - \alpha_0)$, as is obvious from Fig. 85. (If $O$ and $O'$ be the two lattice points, and $O\,P$ and $O'\,P'$ the perpendiculars dropped on the incident and diffracted rays respectively, the difference in path is equal to the difference between the intervals $O\,P'$ and $O'\,P$.) Rays of given wave-length $\lambda$ are therefore diffracted only in those directions (with a cosine $\alpha$), for which the relation

$$a\ (\alpha - \alpha_0) = n\,\lambda \tag{1}$$

is satisfied for *integral* values of the so-called *ordinal number*

[2] $\alpha_0$ is therefore the sine of the angle of incidence.

(*order*) $n$. Conversely, if the incident radiation is composed of rays of all wave-lengths, it is *resolved into spectra* by the diffraction, a different diffraction spectrum corresponding to each value of the ordinal number $n$. (In passing it may be mentioned that, if we turn the diffracted rays about the direction of the lattice in Fig. 85, we obtain the directions of the light reflected by diffraction, though, of course, it is not a case of reflection in the true sense of the word.)

It follows from eqn. (1) that diffraction can never occur, unless

$$a > \lambda. \tag{2}$$

Moreover, $a$ must not be *much* larger than $\lambda$, because otherwise, in the spectra of low order which are of particular importance, the difference $(\alpha - \alpha_0)$ would be too small and the spectral resolution insufficient.[3]

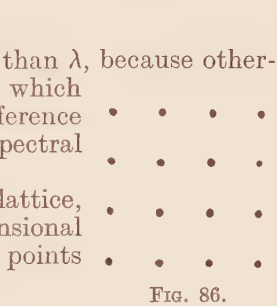


Fig. 86.

From the one-dimensional line-lattice, let us now pass to the two-dimensional *plane-lattice*, which will be formed by points with the coordinates

$$x = k_1 a, \quad y = k_2 b \tag{3}$$

constructed in a plane coordinate system, where $a$ and $b$ are two lattice constants, and $k_1$ and $k_2$ can assume every integral value, either positive or negative. The coordinate axes may be inclined at any angle to each other, but we shall limit ourselves to a rectangular system in which the two lattice constants are identical (Fig. 86).

Let us suppose rays to impinge on such a plane-lattice, making with the $x$- and the $y$-axis angles of which the cosines are $\alpha_0$ and $\beta_0$. Eqn. (1) then expands into a system of two equations, *viz.*,

$$\begin{cases} a\,(\alpha - \alpha_0) = n_1 \lambda\,, \\ a\,(\beta - \beta_0) = n_2 \lambda\,, \end{cases} \tag{4}$$

where $\alpha$ and $\beta$ are the cosines of the angles included between the diffracted ray and the $x$- and the $y$-axis. Since the two equations are mutually independent, they can be solved with respect to the unknowns $\alpha$ and $\beta$. Hence, to every value of the wave-length there corresponds, for a given direction of

[3] In the best optical line-gratings (*Rowland* concave gratings), $a$ is always less than ten times the wave-length.

incidence and given ordinal numbers $n_1$ and $n_2$, a definite direction of diffraction determined by the values of $\alpha$ and $\beta$ [according to eqn. (4) ]. Thus a *plane-lattice* also leads to a resolution into spectra of the incident radiation, this resolution forming a *dual manifold of diffraction spectra*.[4]

We accomplish the transition from the plane-lattice to *space-lattices* by imagining a three-dimensional coordinate system in a space filled with points of which the coordinates are

$$x = k_1 a, \quad y = k_2 b, \quad z = k_3 c ; \tag{5}$$

where $a$, $b$, $c$ are lattice constants, and the numbers $k_1$, $k_2$, $k_3$ assume every integral value, positive and negative. When the coordinate system is rectangular, we term the space-lattice *orthogonal*. In that case, if the three lattice constants $a$, $b$, $c$ are all equal, we define the space-lattice as *regular* or *cubic*.

The following considerations are limited to the special case of a *cubic* space-lattice. Let rays meet such a space-lattice in a direction including with the three coordinate axes angles of which the cosines are $\alpha_0$, $\beta_0$, $\gamma_0$. In place of eqns. (4), we have then the following *triplet of equations* :—

$$\left\{\begin{aligned} a\,(\alpha - \alpha_0) &= n_1 \lambda , \\ a\,(\beta - \beta_0) &= n_2 \lambda , \\ a\,(\gamma - \gamma_0) &= n_3 \lambda . \end{aligned}\right. \tag{6}$$

There is, however, an essential difference between this triplet of equations and eqns. (4). In eqns. (4) we were able to regard as unknowns two magnitudes which completely determine the direction of the diffracted ray for a given direction of incidence, for given wave-length and order ; the number of unknowns was the same as the number of equations. In the present case, however, there are only two unknowns in the three equations (6), if we take again as given the direction of incidence, wave-length, and order. For the three directional cosines $\alpha$, $\beta$, $\gamma$ determining the direction of the diffracted ray are not mutually independent ; two of these cosines are always connected with the third in accordance with the elementary relation

$$\alpha^2 + \beta^2 + \gamma^2 = 1 . \tag{7}$$

Since we cannot regard the ordinal numbers as unknowns inasmuch as they can only assume perfectly definite, per-

[4] Diffraction phenomena of this kind are to be observed when a street lamp is viewed through a thin umbrella.

missible values, *viz.*, integral values (and, indeed, only the lower values require consideration), it appears that, in the three equations, the *wave-length* itself must also be an *unknown* for a given direction of incidence. It is not, indeed, limited to a single value, but it may have any of a *series of possible values* corresponding to the different possible integral values of $n_1$, $n_2$, $n_3$.

These values of the wave-lengths can easily be found from eqns. (6) and (7). We write eqns. (6) in the form

$$(8) \qquad \begin{cases} a = a_0 + n_1 \dfrac{\lambda}{a}, \\ \beta = \beta_0 + n_2 \dfrac{\lambda}{a}, \\ \gamma = \gamma_0 + n_3 \dfrac{\lambda}{a}. \end{cases}$$

Then we square these equations and add them together, taking due account of eqn. (7) and the analogous equation connecting the directional cosines $\alpha_0, \beta_0, \gamma_0$. In this way we find

$$1 = 1 + 2\frac{\lambda}{a}(n_1\alpha_0 + n_2\beta_0 + n_3\gamma_0) + \frac{\lambda^2}{a^2}(n_1^2 + n_2^2 + n_3^2)$$

or,

$$(9) \qquad \lambda = -2a\,\frac{n_1\alpha_0 + n_2\beta_0 + n_3\gamma_0}{n_1^2 + n_2^2 + n_3^2}.$$

In contrast to line- and plane-lattices, a *space-lattice* thus gives *no continuous diffraction spectrum;* its diffractive effect, on the contrary, is to *single out from the incident radiation* those rays of which the *wave-lengths* satisfy eqn. (9) for *integral* values of $n_1$, $n_2$, $n_3$.

Eqns. (6) also enable us readily to calculate the *angle* included between the *diffracted* and the *incident ray*. Denoting this angle by $2\theta$, we have [by § 3 eqn. (14)]

$$(10) \qquad \cos 2\theta = \alpha\alpha_0 + \beta\beta_0 + \gamma\gamma_0.$$

On squaring eqns. (6) and adding, we find [bearing eqn. (7) in mind]

$$a^2(2 - 2\cos 2\theta) = \lambda^2(n_1^2 + n_2^2 + n_3^2).$$

But, in accordance with a well-known trigonometrical formula,

$$\sin\theta = \sqrt{\frac{1-\cos 2\theta}{2}}. \tag{11}$$

Hence,

$$\sin\theta = \frac{\lambda}{2a}\sqrt{n_1^2 + n_2^2 + n_3^2}. \tag{12}$$

From the diffracting lattice point $O$, we cut off an interval $O\,A$ one unit in length in the direction of the incident ray produced, and an equal interval $O\,B$ in the direction of the diffracted ray (Fig. 87). Then we construct the *median plane* $MM'$ between the incident and the diffracted ray, by bisecting the angle $A\,O\,B$ and erecting through the bisector a plane at right angles to the plane $A\,O\,B$. Every point on the median plane must therefore be equidistant from $A$ and $B$.

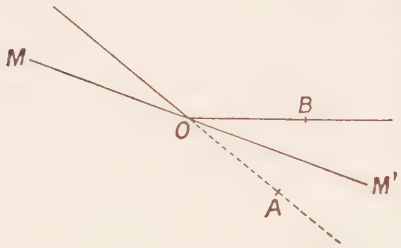


Fig. 87.

With the point $O$ as origin, we construct next a coordinate system with axes parallel to the original axes. With respect to this system, the coordinates of the point $A$ are $\alpha_0, \beta_0, \gamma_0$, while those of the point $B$ are $\alpha, \beta, \gamma$. Since an arbitrary point $(x, y, z)$ on the median plane is equidistant from $A$ and $B$, it must satisfy the relation

$$(x-\alpha_0)^2 + (y-\beta_0)^2 + (z-\gamma_0)^2 = (x-\alpha)^2 + (y-\beta)^2 + (z-\gamma)^2. \tag{13}$$

Bearing in mind eqn. (7) and the analogous one for $\alpha_0, \beta_0, \gamma_0$, we find therefore

$$x(\alpha-\alpha_0) + y(\beta-\beta_0) + z(\gamma-\gamma_0) = 0, \tag{14}$$

or, by eqns. (8),

$$x n_1 + y n_2 + z n_3 = 0. \tag{15}$$

Referred to a lattice point as origin, this is the *equation of the median plane* between the incident ray and the ray diffracted at that point, the order of which is determined by the whole numbers $n_1, n_2, n_3$.

### § 119. Lattice Planes.

In a cubic lattice, let us suppose intervals to be cut off along the three coordinate axes from the origin, these intervals being equal to the quotients of the lattice constant and three whole numbers $k_1$, $k_2$, $k_3$, which have *no common denominator*. The extremities of the three intervals then define a plane, the equation of which we wish to determine.

Since the coordinates of any arbitrary plane are interconnected by a *linear* relation, the most general form of the equation of any *plane* is

$$D_1 x + D_2 y + D_3 z = D_4, \tag{1}$$

where the four magnitudes denoted by $D$ are constants. Let

$$D_4 = D_1 C_1 = D_2 C_2 = D_3 C_3. \tag{2}$$

Eqn. (1) then becomes

$$\frac{x}{C_1} + \frac{y}{C_2} + \frac{z}{C_3} = 1. \tag{3}$$

If now we consider, for instance, the point where the plane cuts the $x$-axis, both $y$ and $z$ are to be equated to zero for this point of intersection, and we see therefore that the three constants $C_1$, $C_2$, $C_3$ simply represent the intercepts of the plane upon the three coordinate axes. The plane which we considered above cuts off an interval of magnitude $a/k$, along the $x$-axis, where $a$ is the lattice constant. Hence the equation of this plane is given by

$$k_1 x + k_2 y + k_3 z = a. \tag{4}$$

In general, we *define as a lattice plane of the space-lattice* every plane for which the left-hand side of eqn. (4) is equal to an *integral multiple of the lattice constant;* as the general equation of a lattice plane, we have therefore

$$k_1 x + k_2 y + k_2 z + s a, \tag{5}$$

where $s$ denotes a whole number. Since, however, $s$ can assume the values of each one of the whole series of integers, the three so-called *indices of the lattice planes* ($k_1$, $k_2$, $k_3$) determine a *set of parallel planes*, one of which, *viz.*, the one for which $s = 0$, passes through the origin.[1] For the rest, $s$ may of course be either positive or negative.

It is also easy to see that an *infinite number of lattice points* will be distributed in a reticular manner over every

[1] For the point ($x = y = z = 0$) satisfies eqn. (5) when $s = 0$.

lattice plane, which is supposed to be unlimited in extent. If $m$ be any integer, positive or negative, and we put

$$\text{(6)} \qquad \begin{cases} x = y = m\,k_3\,a\,, \\ z = -\,m\,(k_1 + k_2)\,a\,, \end{cases}$$

then

$$k_1\,x + k_2\,y + k_3\,z = 0\,.$$

That is to say, all lattice points represented by eqns. (6) lie on the lattice plane ($s = 0$) which passes through the origin. The same result obviously holds good for all lattice points which are represented by the equations to be obtained from eqns. (6) by a cyclic interchange, *viz.*,

$$y = z = m\,k_1\,a\,, \quad x = -\,m\,(k_2 + k_3)\,a\,,$$

or,

$$z = x = m\,k_2\,a\,, \quad y = -\,m\,(k_3 + k_1)\,a\,.$$

Let us now consider any arbitrary plane of the set of lattice planes, for which the value of $s$ differs from zero, and let us form, for instance, the *indeterminate equation*

$$\text{(7)} \qquad k_1\,x' + k_2\,y' = s\,.$$

Naturally this equation always has integral solutions when both $x'$ and $y'$ may be negative: let $x^*$ and $y^*$ be such a pair of solutions. If then we consider the lattice point

$$\text{(8)} \qquad x = x^*\,a\,,\, y = y^*\,a\,,\, z = 0\,,$$

we find for this point, according to eqn. (7),

$$k_1\,x + k_2\,y + k_3\,z = s\,a\,;$$

*i.e.*, the lattice point determined by eqn. (8) belongs to the lattice plane in question. We have therefore only to remove the origin to this lattice point in order to be able to apply the proof advanced above to the lattice plane under consideration, and thus to show that it also contains an infinite number of lattice points.

The *distance between two successive lattice planes* in a set determined by a triplet of indices may be denoted by $d$. This is obviously equal to the *perpendicular* dropped from the origin upon the lattice plane for which $s = 1$. This plane cuts off intercepts $a/k_1$, $a/k_2$, $a/k_3$ along the three coordinate axes. Denoting by $\varphi_1$, $\varphi_2$, $\varphi_3$ the angles included between the perpendicular and the three axes, we have

$$\text{(9)} \qquad d = \frac{a}{k_1}\cos\varphi_1 = \frac{a}{k_2}\cos\varphi_2 = \frac{a}{k_3}\cos\varphi_3\,.$$

But the sum of the squares of the three cosines is equal to unity. Hence it follows from eqn. (9) that

$$d = \frac{a}{\sqrt{k_1^2 + k_2^2 + k_3^2}}. \tag{10}$$

On comparing eqn. (5) with § 118 eqn. (15), we recognize, first, that the *median plane* between the incident ray and the ray diffracted in a space-lattice always represents a *lattice plane*. We see further, however, that the indices of this lattice plane always correspond to the quotients of the ordinal numbers, determining the order of the diffracted ray, and their highest common factor. We have

$$n_1 = n\,k_1\,, \quad n_2 = n\,k_2\,, \quad n_3 = n\,k_3\,, \tag{11}$$

where $n$ is the highest common factor of the ordinal numbers $n_1, n_2, n_3$. The diffraction of the incident rays may therefore be regarded as a *reflection at the lattice planes* of the space-lattice, the simple relation expressed by eqn. (11) connecting the ordinal numbers of the diffraction with the indices of the reflecting lattice planes.

The angle made by the incident ray with the lattice plane, *i.e.*, the complement of the angle of incidence, is called the *glancing angle*. In accordance with eqn. (11) and § 118 eqn. (12), it is determined by the relation

$$\sin\theta = \frac{n\,\lambda}{2\,a}\sqrt{k_1^2 + k_2^2 + k_3^2}. \tag{12}$$

In virtue of eqn. (10), however, this may also be written

$$2\,d\sin\theta = n\,\lambda\,, \tag{13}$$

where $n$ is an arbitrary integer.

If now we consider two lattice planes, the distance ($A\ B$) between which is equal to $d$ (Fig. 88), the difference in path between two rays reflected at the points $A$ and $B$ will be equal to ($C\ B + B\ D$), where $C$ and $D$ are the feet of the perpendiculars dropped from the point $A$ on the directions of the ray reflected by the second plane. But

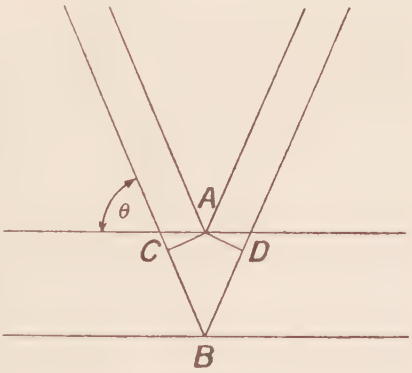


FIG. 88.

$(C\ B + B\ D)$ is equal to $2\ d \sin \theta$. Hence, a set of lattice planes *singles out* from the incident radiation those rays for which an *integral multiple of the wave-length* is *equal to the difference in path* between two rays, which are reflected from two successive lattice planes.

As a matter of fact, this law [which is, of course, only another way of expressing § 118 eqn. (9) ] is very plausible. For, if the diffraction of the rays takes place in such a way that they appear to be reflected at the lattice planes in accordance with the usual laws of optics, then the reflected rays must strengthen and weaken each other owing to interference. Moreover, an appreciable increase in intensity will be observed only when the difference in path between the rays reflected at two successive planes amounts exactly to an integral multiple of a wave-length.

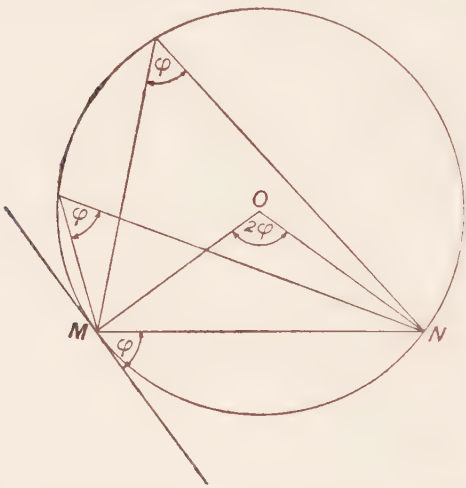


Fig. 89.

Furthermore, as we have shown in the previous considerations from two entirely different points of view, a space-lattice only diffracts incident rays of quite definite wave-length from radiation with a given direction of incidence. Thus complete diffraction of a narrow bundle of parallel rays can obviously be attained only if the *space-lattice be rotated*. A complete *resolution into spectra* is then actually attainable, as may be seen from some simple geometrical relations.

In order to derive these relations, we start from a well-known elementary theorem, which states that all the *peripheral angles subtended by the same chord in a circle* are *equal in magnitude*, and also equal to the *tangential angle* included between the chord and the tangent to the circle, drawn at one end of the chord (Fig. 89).[2] We shall consider (Fig.

[2] The second part of the theorem, which is also well known, is most simply proved from the fact that the angle at the centre is twice the angle at the

90) a *lattice plane*, and suppose a narrow bundle of rays, passing through a slit $R$, to meet it at the point $O$ with a glancing angle $\theta$. Denoting the position of the plane by the straight line $A\,A'$, we have therefore

$$\angle\, R\,O\,A' = \theta\,. \tag{14}$$

We draw a circle with $R\,O$ as a chord, and the line $A\,A'$ touching it tangentially at $O$.

Let us now *rotate* the space-lattice about an axis perpendicular to the plane of the figure, so that the lattice plane turns from the position $A\,A'$ into the position $B\,B'$. Let the straight line $B\,B'$ intersect the circle at the point $P$. Then the angle $R\,P\,O$, a peripheral angle subtended by the chord $R\,O$, is, according to the theorem just quoted, equal to the tangential angle $R\,O\,A'$. Thus we have also

$$\angle\, R\,P\,O = \theta\,. \tag{15}$$

By eqn. (13), however, a definite value of the glancing angle is associated with a definite value of the wave-length, for a given ordinal number $n$.

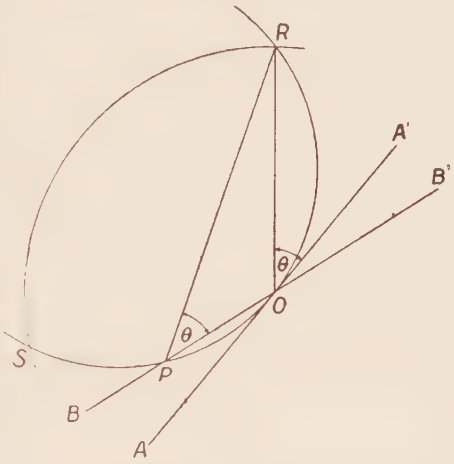



Fig. 90.

When the space-lattice is *rotated*, the point in the lattice plane at which the reflection of a ray of definite wave-length takes place thus *travels along the circumference of the circle* which, passing through the points $R$ and $O$, touches at $O$ the straight line $A\,A'$ that makes an angle $\theta$ with the line $R\,O$. While the lattice plane is being rotated, the ray of definite wave-length therefore *glides* along it.

We suppose now (Fig. 90) a *second circle* to be described about $O$ as centre, with $O\,R$ as radius. Let this second circle cut the first circle $R\,O\,P$ at a point $S$. It may be easily shown that every ray which strikes the lattice plane at a

circumference. Hence in Fig. 89 the angle $M\,O\,N$ is equal to $2\phi$, so that the angle $O\,M\,N$ is equal to half of $(180 - 2\phi)$. The tangential angle, however, is complementary to the angle $O\,M\,N$, because the radius is perpendicular to the tangent. Thus the tangential angle is equal to $\phi$.

glancing angle $\theta$ must, after reflection, pass through the point $S$, whatever may be the position of the lattice plane at the instant of reflection.

We see (Fig. 91) that the angle $O\,S\,R$, a peripheral angle subtended by the chord $O\,R$, is equal to the angle $R\,P\,O$; hence, by eqn. (15),

$$\angle\, O\,S\,R = \theta. \tag{16}$$

But $S\,O\,R$ is an isosceles triangle, and it follows therefore from eqn. (16) that the angle at the apex of this triangle

$$\angle\, S\,O\,R = 180^\circ - 2\,\theta. \tag{17}$$

Thus, by eqn. (14),

$$\angle\, S\,O\,A = \theta. \tag{18}$$

*i.e.*, the ray reflected at the point $O$, when the lattice plane is in the position $A\,A'$, passes through the point $S$ after reflection.

Moreover, the angles $S\,O\,R$ and $S\,P\,R$ are peripheral angles subtended by the same chord $S\,R$. Hence by eqn. (17) we have

$$\angle\, S\,P\,R = 180^\circ - 2\,\theta, \tag{19}$$

so that, by eqn. (15),

$$\angle\, S\,P\,B = \theta. \tag{20}$$

Therefore, the ray reflected at the point $P$, when the lattice plane is in the position $B\,B'$, likewise passes through the point $S$ after reflection. Furthermore, since the position $B\,B'$ is an entirely *arbitrary* position of the lattice plane, every ray proceeding from the point $R$, which is reflected at an angle $\theta$ from the lattice plane containing $O$, passes through the point $S$ after reflection. For a given order of diffraction, however, a quite definite value of the wave-length corresponds to each glancing angle. Hence, on rotation of the space-lattice, all

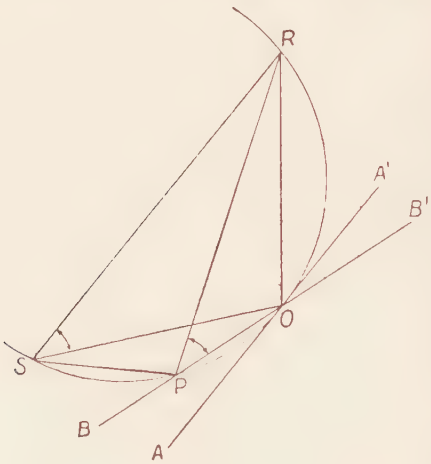



Fig. 91.

rays of the same wave-length are, as it were, "*focussed*" at one and the same point.

### § 120. The Interference of X-Rays in Crystals.

If *crystals* possess the structure of *space-lattices*, and their lattice constants, though larger, are not much larger than the wave-lengths of *X-rays*, these rays should be *diffracted by interference* within the crystals. This diffraction will follow the laws derived from purely geometrical considerations in the previous sections, for the *atoms* corresponding to the *lattice points* must obviously act as *diffraction centres* with respect to electromagnetic waves. An *experimental proof of the interference of X-rays* produced by *crystals* would therefore, on the one hand, demonstrate the lattice structure of crystals, and, on the other hand, confirm the previously expressed conjecture that X-rays are electromagnetic waves with a wave-length of the order of $10^{-9}$ cm.

Starting out with this idea, *Laue*, in 1912, was the first to investigate the action of crystals on X-rays, and actually to demonstrate experimentally the suspected phenomena. In these researches, which were carried out by Laue in collaboration with *Friedrich* and *Knipping*, a narrow parallel pencil of rays was separated from the radiation of an X-ray tube by means of lead screens with adjustable slits. The rays fell on a plate of crystal about 1 cm. square and $\frac{1}{2}$ mm. thick, and a photographic plate was mounted some centimetres behind the crystal.

Now, from a parallel bundle of incident rays, a stationary space-lattice diffracts rays only in quite definite directions, which, apart from the direction of incidence, depend only on the ordinal numbers of the diffraction. [These directions are defined by § 118 eqn. (8), in which $\lambda$ is to be given the value following from § 118 eqn. (9).] The points at which the directions of these diffracted rays *meet* the plane of the photographic plate must thus be *blackened* on the plate.

Experiment actually gave these results. In Figs. 92 and 93 (on Plate II) are reproduced the so-called *Laue photographs* which were obtained with a crystal of *zinc blende* (ZnS). In the exposure shown in Fig. 92, the distance of the plate from the crystal amounted to 3·5 cm., whereas it was only 1 cm. in the exposure reproduced in Fig. 93. On that account, the spots appear more closely packed in the second photograph.

In both exposures, the rays met a *face of the cube at right angles*. Their direction was therefore that of an edge of the cube, and this explains the *fourfold symmetry* to be observed in the photographs. An axis of symmetry in a crystal is termed $n$-fold, when through this axis $n$ planes can be drawn, with respect to which the crystal is symmetrical. Now it is easy to see that each *edge of a cube* in a regular crystal represents a *fourfold axis of symmetry*. For, through each edge, four planes of symmetry can be drawn, their lines of intersection with a plane normal to the edge being represented in Fig. 94; two of these intersecting lines are again edges, and two face diagonals. Thus the photographs exhibit a fourfold symmetry, provided that the incident pencil of rays be perpendicular to a face of the cube. Corresponding to Fig. 94, four lines of symmetry are to be recognized, and hence all spots (provided they do not lie just on a line of symmetry) are repeated *eight times* in all. This is actually to be seen in Figs. 92 and 93. If the ray falls perpendicularly on an *octohedral* face cut from a regular crystal, geometrical considerations (into which we cannot enter here) show that its direction is that of a trigonal axis of the cube—a threefold axis of symmetry. An exposure made in this way is reproduced in Fig. 95 (on Plate II), in which each spot occurs *six times*.



Fig. 94.

In 1913, *W. H.* and *W. L. Bragg* (father and son) deduced from the Laue photographs most important conclusions on the *structure of crystals*. Although the space-lattice nature of crystals had been proved beyond doubt, yet, before crystals could be actually utilized for accurate measurements of wave-lengths, the exact lattice structure of at least *one* crystal had to be determined, and in this connection there were two questions to be solved.

The first question concerns a possible *centring of the space-lattice*. A cubic lattice may be simple, space-centred, or finally, face-centred. The symmetry relations are the same in all three cases. A cubic lattice (such as was considered in §§ 118 and 119) is *space-centred* when, in all the elementary cubes, the edges of which are given by the lattice constant, an additional lattice point is situated at the centre of the cube; each elementary cube then contains nine instead of eight lattice points. If, on the other hand, the interior of the cube is unoccupied, whilst lattice points are



FIG. 92.

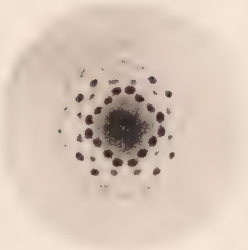


FIG. 93.



FIG. 95.
Laue Photographs.

[*To face p.* 166.

situated at the centres of the six faces (Fig. 96), a *face-centred* lattice results.

The second fundamental question of crystal structure arises from the fact that most crystals are composed of *different kinds of atoms, e.g.*, rock-salt is made up of sodium and chlorine atoms. It might well be expected that each kind of atom would form a separate space-lattice, but the question was still unsolved as to how the different space-lattices were combined in the crystal.



Fig. 96.

Very ingenious reasoning, which we cannot discuss here,[1] enabled W. H. and W. L. Bragg to answer these questions for numerous crystals, by investigating the Laue photographs. The arrangement and number of the interference spots differ, according as to whether and how the space-lattice is centred. Moreover, the relative positions of the combined space-lattices of the various kinds of atoms could be determined from the fact that the behaviour of an atom with respect to X-rays depends on the number of electrons contained in the atom, and thus, broadly speaking, on the *atomic weight*. Hence a lattice plane of atoms reflects X-rays in a different manner according as the two species of atom have much the same, or widely differing atomic weights. Thus, in spite of a presumably identical crystal structure, the different alkali halides, for example, should give rather different Laue photographs. [*E.g.*, in potassium chloride (sylvine) the two atomic weights only differ slightly (39 and 35·5), but in potassium bromide they are very different (39 and 80).] A comparison of the photographs of various alkali halides thus rendered possible a recognition of the structure of crystals of these compounds. In a later section we shall discuss the general results of crystal analysis; for the moment, we shall refer to the structure only of the alkali halides, because one of them, rock-salt (NaCl), is the principal crystal used in the measurement of the wave-lengths of X-rays.

In a *crystal of rock-salt*, the atoms form a *simple cubic lattice* in the sense of the considerations of the earlier sections (§§ 118 and 119), if we disregard the difference between the Na and Cl atoms. For the Na atoms by themselves form a face-centred lattice, as do the Cl atoms likewise, and the two lattices are combined in such a way that the points of the one lattice lie at the mid-points of the edges of the other

[1] *Cf.*, for instance, Sir W. H. and Prof. W. L. Bragg's "*X-Rays and Crystal Structure*" (London: Bell, 1924).

lattice (*cf.* Fig. 97, where the two kinds of atoms are distinguished by black and by white circles).

Now once the structure of rock-salt crystals is known, the *distance between the lattice planes* can immediately be calculated, taking the faces of the cubes as lattice planes. Denoting this distance by $d$, we see that each cube, the length of an edge of which is $d$, contains a total of eight atoms, which are situated at its corners. Since, however, each corner belongs to the eight cubes which meet at it, the number of atoms contained in a cubic centimetre is given by

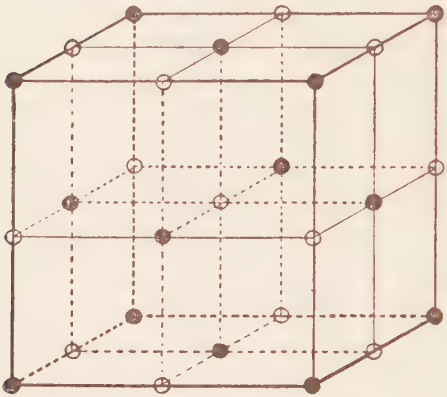


FIG. 97.—Structure of a rock-salt crystal.

$$Z = \frac{1}{d^3}. \tag{1}$$

On the other hand, if $M$ be the so-called *molecular weight* of rock-salt, $\varrho$ its *density*, and $L$ the *Loschmidt number*,

$$\varrho = \frac{M}{L} \cdot \frac{Z}{2}. \tag{2}$$

We have to divide by two, because it is the mass of *two* atoms (one Na and one Cl atom) taken together that is equal to the molecular weight divided by the Loschmidt number. Combining eqns. (1) and (2), we find the distance between the lattice planes to be given by [2]

$$d = \sqrt[3]{\frac{M}{2\,L\,\varrho}}. \tag{3}$$

The molecular weight of NaCl is equal to the sum of the atomic weights of Na (23·00) and Cl (35·46), *i.e.*,

$$M = 58{\cdot}46. \tag{4}$$

[2] In other lattice structures the numerical factor in eqn. (3) would, of course, be different.

The value found for the density of rock-salt from very accurate measurements is

$$\varrho = 2{\cdot}164\,. \tag{5}$$

Hence, by eqn. (3) and § 87 eqn. (6),

$$d = 2{\cdot}814 \times 10^{-8}\ \text{cm.} \tag{6}$$

This value has been derived with an accuracy of about 1/3%, depending chiefly on the uncertainty occasioned by the Loschmidt number. Since, however, X-ray spectroscopic measurements can be carried out with a much greater accuracy, an *ideal value* of $2{\cdot}81400 \times 10^{-8}$ cm. is assigned to $d$ for such purposes.

As soon as a fundamental constant in the theory of crystals is thus evaluated, very exact *measurements of the wave-lengths* of X-rays at once become possible. These may be most simply carried out by means of the *method of the rotating crystal*, devised by *W. H.* and *W. L. Bragg* in 1913. This method is based, on the one hand, on the law of reflection at the lattice planes derived in the previous section and likewise due to the Braggs, and, on the other hand, on the fact that, as far as the *reflection of X-rays by crystals* is concerned, practically only the set of lattice planes *parallel to the front face of the crystal* requires consideration.[3]

Accordingly, a very narrow pencil of X-rays from a small lead diaphragm is directed upon a crystal, which is mounted on a goniometer table, rotating about a vertical axis, in such a way that the axis of rotation lies in the front face of the crystal.[4] The crystal is very slowly rotated by hand, or by clockwork. According to § 119 eqn. (13), rays of the wave-length $\lambda$ are then reflected only for quite definite glancing angles, *viz.*, only for those angles which, for integral (and low) values of the ordinal number $n$, satisfy the equation

$$2\,d \sin\theta = n\,\lambda\,. \tag{7}$$

As the crystal rotates, the ray glides along its face; the advantage of this method over one where the crystal is kept in the same position is that the disturbing effects of the unavoidable defects in the crystal face are thereby eliminated.

Now it follows from the geometrical considerations of § 119 that rays of the same wave-length are *focussed* at points

[3] Naturally, the face itself is always a lattice plane.

[4] In *Seeman's* apparatus a sharp edge is fixed directly against a part of the crystal face that is as perfect as possible. Edge and face together thus form a fine slit.

on the circumference of a circle which passes through the slit, while its centre lies on the axis of rotation. If, therefore, a *photographic film* be placed round this circle, and if the radiation contains only rays of definite wave-lengths, *line spectra* of several orders must be formed on the film. The spectrum of the first order can at once be recognized, for it is far more intense than that of the second order, while the latter is more intense than the spectrum of the third order, and so on.

In general, only the *spectrum of the first order* is utilized. The wave-lengths of the lines contained in it are readily found by constructing a *scale* of wave-lengths in accordance with eqn. (7) (wherein we put $n = 1$). The crystal of rock-salt may, of course, be replaced by any other, quite arbitrary crystal, even of unknown structure, provided only that one line can be identified with certainty. A comparison of the positions of this line in the scales of the two crystals immediately enables the scale of the crystal used to be converted into that of rock-salt.

By means of the photographic method, a spectroscopic accuracy of measurement of almost one-millionth part of the measured value can be obtained. With this accuracy the wave-lengths of *different* lines can be compared, and, as was mentioned above, this necessitates the use of an ideal value for the rock-salt constant, which cannot be so accurately determined. The *photographic* method is due to *Maurice de Broglie* (1913); W. H. and W. L. Bragg were then still working with a so-called *ionization spectrometer*. This instrument consists of a lead cylinder filled with gas, and provided with a narrow slit through which X-rays can pass into the gas. The cylinder is rotated around the crystal along the circle on which rays of equal wave-length are focussed, in such a way that the window always remains on the circle. Rays entering the window ionize the gas, and its resulting conductivity can be accurately measured. More recently the ionization method has been very much improved, and it is principally used for measurements of intensity in X-ray spectra because the intensity of the X-rays is approximately proportional to the exactly measurable ionization. It may also be remarked that wave-lengths can be determined by the perfected ionization method with an accuracy of one in ten thousand.

If now we consider a small region of the spectrum extending from a wave-length $\lambda$ to a wave-length $(\lambda + \Delta\lambda)$, the *spectral resolution* of this region is defined by the difference

quotient $\Delta\theta/\Delta\lambda$, and hence approximately by the differential coefficient. If follows from eqn. (7) that

$$\frac{d\theta}{d\lambda} = \frac{n}{2\,d\cos\theta}. \tag{8}$$

The resolving power thus increases with the ordinal number, but this advantage is counterbalanced by the rapid decrease in intensity with increasing ordinal number. Nevertheless, close doublets are often measured in the second order spectra, sometimes even in the third, fourth, or fifth orders. Moreover, as eqn. (8) shows, the resolution increases with the glancing angle $\theta$, and it becomes especially strong when the angle $\theta$ approaches the value of 90°, so that sin $\theta$ is almost equal to unity. According to eqn. (7), however, this is the case when twice the distance between the lattice planes becomes nearly equal to an integral multiple of the wave-length of the region to be resolved. The resolution of a given small region, such as a doublet, is therefore best effected with a crystal which satisfies this relation for as small an ordinal number $n$ as possible.

Since neither sin $\theta$ can be greater, nor $n$ smaller, than unity, it follows from eqn. (7) that the greatest wave-length which can ever be diffracted by a crystal is equal to twice the distance between the lattice planes, which in a rock-salt crystal is approximately equal to $5{\cdot}6 \times 10^{-8}$ cm.[5] The unit of wave-length used in X-ray spectroscopy is the so-called *X-unit* (X), which is equal to one-thousandth of the optical spectroscopic unit, the Ångström unit. Hence

$$1\text{ X} = 0{\cdot}001\text{ Å.} = 10^{-11}\text{ cm.} \tag{9}$$

Thus with a rock-salt crystal, only wave-lengths up to about 5,000 X can be measured. For the measurement of larger wave-lengths, other crystals with larger distances between the lattice planes are used, such as selenite or potassium ferrocyanide, in which the lattice plane distances are $7{\cdot}6 \times 10^{-8}$ and $8{\cdot}5 \times 10^{-8}$ cm. respectively. By means of these, wave-lengths as large as 13,000 X (= 13 Å.) can be measured, whilst, with the help of *crystals of solid fatty acids*, an extension to about 50,000 X is possible. Apart from rock-salt, the principal crystal used in wave-length determinations is *calcite*, in which the distance between the lattice planes is $3{\cdot}029 \times 10^{-8}$ cm. It may also be observed that soft X-rays (*i.e.*, rays of large wave-lengths) are very

[5] If the face of a regular crystal is not that of a cube, the lattice plane distance is less than $d$.

strongly absorbed by gases even at quite low pressure, so that, in the determination of such wave-lengths, the whole path of the rays has to be *in vacuo*.

With very *hard* rays the method of the rotating crystal cannot be used. Hard rays penetrate too deeply into the crystal and thereby produce a strong broadening effect on the lines. Hence, for very hard rays another method is used, which was developed by *Rutherford* and *Andrade*, in order to measure the wave-lengths of the extremely hard $\gamma$-rays. A widely diverging beam of rays falls on a thin plate of crystal, and the rays are reflected (at least, in a regular crystal) principally by the lattice planes that are normal to the crystal face struck by the rays. The crystal is fixed in one position, and behind it is placed a slit, which singles out from the emerging beam rays reflected at a certain glancing angle, and hence of a quite definite wave-length. By means of this method, into the details of which we cannot enter here, wave-lengths as small as 50 X have been measured.

Finally, an important method of X-ray analysis that we shall mention is the *crystal powder* method, developed on the one hand by *Debye* and *Scherrer*, and on the other hand by *Hull*. In this form, the time succession of diffraction centres in the rotating crystal method is replaced by a corresponding succession in space. The powder, in fact, consists of numerous minute crystals mingled together in all possible orientations. Hence, when a narrow pencil of rays strikes a small tube filled with the powder, reflection occurs at every possible angle, so that in this way also, the incident radiation is resolved into spectra. This method is of great importance for the purpose of crystal analysis.

### § 121. The Relation between Electron Rays and X-Rays.

If the concept of *material waves* is correct, *phenomena similar* to those observed with *X-rays* should also appear with *electron rays*, the equivalent potential for which lies between about 10 and 100,000 volts. For these potentials are associated with velocities of between $10^8$ and $10^{10}$ cm./sec., and hence with characteristic wave-lengths of between $10^{-7}$ and $10^{-9}$ cm.

In 1927 *Davisson* and *Germer*[1] actually succeeded in proving that when electron rays impinge on *crystals*, interference may be produced exactly as in the case of X-rays.

[1] C. J. Davisson and L. H. Germer, *Phys. Rev.*, **30** (1927), p. 705, and *Nature*, **119** (1927), p. 558; for a more detailed report *vide* the *Bell System Technical Journal*, **7** (1928), p. 1.

Davisson and Germer directed a narrow pencil of electrons at right angles to a small slice cut from a monocrystal of nickel, and, with the help of a galvanometer, they investigated the intensity of scattering of the electrons in its dependence on the direction. They were able to detect sharply developed *maxima* for certain directions, and also to show that these directions *coincided* with those in which X-rays would be reflected by the crystal, provided that the known structure of a nickel crystal was assumed to be *contracted*, in a direction parallel to that of the primary electron rays, in a ratio lying between 0·7 : 1 and 0·9 : 1. A reason for the existence of this so-called contraction factor appears in the *work* which has to be done by the electrons on their *emergence from the crystal*. Davisson and Germer were also able to compute the characteristic *wave-lengths* of the electron rays from the diffraction phenomena, and, in agreement with de Broglie's theory, they found that these wave-lengths were equal to the elementary quantum of action divided by the product of the electronic mass and velocity.[2]

An important supplement to Davisson and Germer's researches has been furnished by *G. P. Thomson's* [3] experiments on the penetration of very thin *foils* of gold and other metals by very rapid electrons, with potentials ranging from 20,000 to 30,000 volts. Thomson's photographs (*vide* Fig. 98 on Plate III) showed the same peculiar *diffraction rings* as were obtained long before by Debye and Scherrer, when they subjected very hard X-rays of about the same wave-length (approximately $10^{-9}$ cm.) to diffraction by thin metal foils. He was also able to calculate the lattice constants of the metals employed from the photographs taken with cathode rays, and he obtained, in fact, the same values as had previously been found from X-ray photographs.

Whereas Thomson's electron rays corresponded to extremely hard X-rays, *Rupp*,[4] using exceedingly thin metal foils, obtained some beautiful diffraction phenomena with comparatively *slow electrons* of 150–320 volts; these rays therefore correspond to very *soft* X-rays. Rupp's

[2] It is interesting to note that, in 1923, Davisson and Kunsman had observed that slow electrons are selectively reflected by metallic surfaces [*Phys. Rev.*, **22** (1923), p. 243; *vide* also the summary of these and cognate experiments, executed in the research laboratories of the Bell Telephone Co., in New York, which appears in K. K. Darrow's "*Introduction to Contemporary Physics*" (New York: van Nostrand Co., 1926), ch. v.]

[3] G. P. Thomson, *Proc. Roy. Soc. A*, **117** (1928), p. 600.

[4] E. Rupp, *Ann. der Phys.*, **85** (1928), p. 981.

experiments were based on a method devised by *Cole*,[5] consisting in the blackening of ordinary photographic plates by slow electrons, and on the almost simultaneous discovery by *Becker* [6] that metal foils evince a *selective permeability* for electrons of a certain range of velocities. Quite recently Rupp [7] has also succeeded in producing *electron diffractions with optical gratings*, and from these he has calculated the de Broglie wave-lengths with an accuracy of about 5%.

The concept of diffractible material waves, moreover, has provided an explanation for the formerly enigmatical observations on the *passage of free electrons through gases*. Thus *Ramsauer* [8] discovered that, in the rare gases,[9] the mean free path of electrons increases with decreasing velocity in such a way that the *slow electrons* apparently pass, *almost unchecked*, through the atoms. *Elsasser* [10] was the first to refer the so-called *Ramsauer effect* to a diffraction of material waves, and this explanation was satisfactorily confirmed by *Dymond's* [11] observation that, in the *scattering of electrons* on their passage through *helium*, the curve expressing the *relation between intensity and angle of scattering* exhibits strongly marked *maxima*.[12]

[5] K. Cole, *Phys. Rev.*, **28** (1926), p. 781.

[6] A. Becker, *Ann. der Phys.*, **84** (1927), p. 778.

[7] E. Rupp. *Naturwiss.*, **16** (1928), p. 656.

[8] *Cf.* the summary by R. Minkowski and H. Sponer in "*Ergebnisse der exakten Naturwissenschaften*" (Berlin: Springer, 1924), Vol. III.

[9] Since the rare gases are chemically passive, their atoms have no affinity for electrons.

[10] W. Elsasser, *Naturwiss.*, **13** (1925), p. 711. *Einstein* had previously endeavoured to interpret the phenomena of *gas degeneration* in the light of de Broglie's theory.

[11] G. Dymond, *Nature*, **118** (1926), p. 336; *Phys. Rev.*, **29** (1927), p. 433.

[12] In this connection, attention may be drawn to an observation by *P. Günther*, who, on Nernst's suggestion, investigated the viscosity (internal friction) of hydrogen at very low temperatures, and thereby discovered that, on reaching a critical zone of temperature, the viscosity suddenly diminishes at an increased rate. *Vide Zeitschr. f. physikal. Chem.*, **110** (1924), p. 626.

## Chapter XV

# THE ATOMIC NUCLEUS

### § 122. Isotopy.

As we have seen in the previous considerations, the position of an element in the periodic system depends solely on its *nuclear charge*. Furthermore, we must assume the *nuclei* of the elements to be *built up* of protons and electrons. Hence, provided only that the *resultant nuclear charge*—the *algebraical sum* of all the positive and negative elementary charges forming the nucleus—be the same in two atoms, these must exhibit the same chemical behaviour, even though the number of protons and nuclear electrons differs in the two. Whereas, however, the nuclear charge, and thus the *chemical character*, appears to be determined by the *difference* in the numbers of protons and electrons in the nucleus, the *mass* of the atom, and therefore the so-called *atomic weight*, depends *only* on the number of *protons*, since the mass of an electron can be neglected in comparison with that of a proton. It is thus quite possible for two substances despite *different atomic weights* and *different nuclear structures*, to agree in their chemical and in many important physical properties to such an extent that they would really represent only *different species of one and the same element*.

The first phenomenon of this kind to attract the notice of physicists was the *inseparability* of certain *radioactive elements*. When, for instance, salts of *thorium* and *ionium* (the parent substance of radium) were in any way mixed together, it was found quite impossible to separate them again by any method whatsoever; moreover, the *spectra* of thorium and ionium were observed to be quite identical. Similarly, radium and mesothorium 1 were found to be inseparable, as also were lead and radium D (*cf.* § 126).

All these cases proved that substances differing in their radioactive behaviour could yet be identical in chemical and also in physical respects. Recognition of this fact led *Soddy*, in 1910, to the belief that also the *inactive* (non-radioactive) *elements* might represent *mixtures of several, chemically inseparable species*. Soddy further pointed out that

this assumption would solve a great difficulty, which still faced the atomic theory at that time. This difficulty arose from the *deviations* of the *atomic weights from whole numbers*.

If atoms are actually built up of protons (*i.e.*, hydrogen nuclei) and electrons, and if the mass of the electrons can be neglected in comparison with that of the protons, then all atomic weight numbers must obviously be *integral*. Now, although this is the case with many elements, and especially with the most abundant ones, yet in other elements such as chlorine, the deviations from whole-number values are so considerable that they could not at first be reconciled with that necessary consequence of atomic theory. This difficulty, however, vanishes if, according to Soddy's assumption, the elements really represent *mixed elements*. For, in that case, only the significance of an average weight would attach to the atomic weight, and its constancy would be explained on the grounds that any mixture ratio, once formed (*e.g.*, in times when the earth was in a gaseous state), must remain unaltered for ever, in view of the inseparability of the mixed atoms.

This conception, suggested by Soddy in 1910, was experimentally confirmed in 1913 by *J. J. Thomson*, who showed by means of *positive ray analysis* that the rare gas *neon* really consists of *two atomic species* of *different atomic weights* (20 and 22). Accurate experiments made in 1914 also confirmed the opinion, expressed by Soddy, that the *atomic weight of lead* must differ in value, according as to whether we are dealing with "ordinary" lead, or with lead which has been obtained from uranium minerals or from thorium minerals. (As we shall show in § 126, both uranium and thorium are gradually transformed into lead.) Whereas the atomic weight of ordinary lead is found to be 207·2, atomic weights as low as 206·05 have actually been determined for lead from uranium minerals, and as high as 207·9 for lead from thorium minerals.[1] The term *isotopes*, suggested by Soddy, soon came into general use to denote substances which have different atomic weights, but exhibit the same chemical properties.[2]

Researches on isotopes, properly speaking, began in 1919 with *Aston's* development of positive ray analysis into the so-called *mass-spectroscopy*, by means of an artifice already

[1] *Cf.* O. Hönigschmid, *Zeitschr. f. Electrochemie*, **24** (1918), p. 163, and **25** (1919), p. 91.

[2] *Isos topos* in Greek means "the same place." The name was suggested by the fact that isotopes occupy the same place in the periodic system of the elements.



Fig. 98.—Diffraction of cathode rays on passage through gold foil (left) and celluloid foil (right)—after G. P. Thomson.

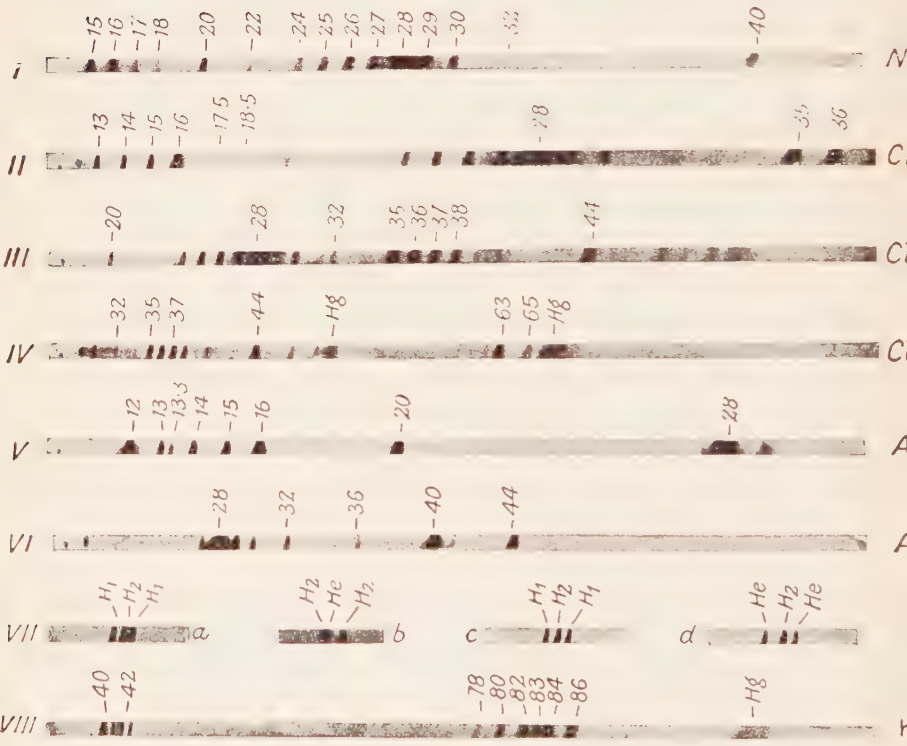



Fig. 99.—Mass-spectra (after Aston).

[To face p. 1

described (*cf.* § 89). Using his new method, Aston was first able to prove, with greater accuracy, that the rare gas *neon* does actually consist of *two isotopes* with the exactly *integral atomic weights* 20·0 and 22·0. Shortly afterwards he was also able to prove that *chlorine*, in which the deviation from a whole-number value had always been felt to be particularly embarrassing, is a mixture of two isotopes with the exactly integral atomic weights 35·0 and 37·0 : it follows from the mixture weight (35·46) that the ratio of the lighter to the heavier gas is about three to one, as can also be recognized from the intensity of the lines in the mass-spectra. Aston has succeeded in gradually increasing the accuracy of the determination of mass to 1 in 10,000.[3] In Fig. 99 (Plate III), Aston's *mass-spectra* of neon, chlorine, argon, and krypton are reproduced, showing most clearly the integral nature of the pure atomic weights.[4]

A summary of the results of researches (up to the end of 1928) on the subject of isotopes is contained in Table XIX ; these researches have been carried out, apart from by Aston himself, principally by *G. P. Thomson* and *Dempster*. Table XIX is supplemented by Table XX, referring to the ten radioactive elements which, together with two gaps, occupy the last twelve places in the periodic system.

As regards the *number of isotopes*, an evident distinction exists between elements of *even* and those of *odd atomic number*. The latter possess comparatively few isotopes, either being so-called pure elements (with only one atomic species) or else exhibiting *two isotopes*, the mass-numbers of which (with the sole exceptions of lithium and boron at the beginning of the series) are *both odd* and *differ by two*. More than two isotopes have been found only in elements of even

[3] F. W. Aston, *Proc. Roy. Soc. A.*, **115** (1927), p. 487.

[4] Lines of known atoms or groups of atoms serve as reference lines in these spectrograms, especially the lines of carbon (12) ; the C-complexes, CH, $CH_2$, $CH_3$, $CH_4$ (13, 14, 15, 16) ; the $C_2$-complexes (24, 25, 26, 27, 28, 29, 30) ; O (16), OH (17), $OH_2$ (18), CO (28), and $CO_2$ (44). Spectrum I shows, besides these lines, which almost always occur, the two neon lines very clearly at 20 and 22. In spectra II, III, and IV can be seen the lines of $Cl_{35}$ and $Cl_{37}$, as well as lines at 36 and 38 due to $Cl_{35}H$ and $Cl_{37}H$. We may also notice " second order lines " at 17·5 and 18·5. These are produced by chlorine ions with double charges, so that the specific charge is then twice as large, which would also be the case if the charge remained single and the mass were halved. Lines 63 and 65 are due to $COCl_{35}$ and $COCl_{37}$. Spectra V and VI show not only the first order lines of argon, but also, quite clearly, a second order line at 20 and a third order line at 13·3. Spectrum VIII is that of krypton, showing plainly five strong lines at 80, 82, 83, 84, 86 and a sixth, weaker, line at 78. These lines also occur in the second order in the mass-spectrum of krypton. An explanation of spectrum VII may well be omitted here.

## Table XIX

### *Atomic Species*

(The alphabetical affixes in the last column indicate the relative shares of the isotopes in the mixed element, *a* denoting the strongest component, *b* the next strongest, *etc.* Where the experimental results are not yet certain, the numbers are enclosed in parentheses.)

| Atomic Number. | Element. | Atomic Weight. | No. of Atomic Species. | Masses of Isotopes. |
|---|---|---|---|---|
| 1 | Hydrogen . | 1·008 | 1 | 1·008 |
| 2 | Helium . | 4·002 | 1 | 4 |
| 3 | Lithium . | 6·94 | 2 | 6*b*, 7*a* |
| 4 | Beryllium . | 9·02 | 1 | 9 |
| 5 | Boron . . | 10·82 | 2 | 10*b*, 11*a* |
| 6 | Carbon . . | 12·000 | 1 | 12 |
| 7 | Nitrogen . | 14·008 | 1 | 14 |
| 8 | Oxygen . | 16·000 | 3 | 16*a*, 17*c*, 18*b* |
| 9 | Fluorine . | 19·00 | 1 | 19 |
| 10 | Neon . . | 20·18 | 2 (3) | 20*a*, (21), 22*b* |
| 11 | Sodium . . | $22·99_7$ | 1 | 23 |
| 12 | Magnesium . | 24·32 | 3 | 24*a*, 25*b*, 26*c* |
| 13 | Aluminium . | 26·97 | 1 | 27 |
| 14 | Silicon . . | 28·06 | 3 | 28*a*, 29*b*, 30*c* |
| 15 | Phosphorus . | 31·02 | 1 | 31 |
| 16 | Sulphur . | 32·06 | 3 | 32*a*, 33*c*, 34*b* |
| 17 | Chlorine . | 35·46 | 2 | 35*a*, 37*b* |
| 18 | Argon . . | 39·94 | 2 | 36*b*, 40*a* |
| 19 | Potassium . | $39·10_4$ | 2 | 39*a*, 41*b* |
| 20 | Calcium . | 40·07 | 2 | 40*a*, 44*b* |
| 21 | Scandium . | 45·10 | 1 | 45 |
| 22 | Titanium . | 47·90 | 1 (2) | 48, (50) |
| 23 | Vanadium . | 50·95 | 1 | 51 |
| 24 | Chromium . | 52·01 | 1 | 52 |
| 25 | Manganese . | 54·93 | 1 | 55 |
| 26 | Iron . . | 55·84 | 2 | 54*b*, 56*a* |
| 27 | Cobalt . . | 58·94 | 1 | 59 |
| 28 | Nickel . . | 58·69 | 2 | 58*a*, 60*b* |
| 29 | Copper . . | 63·57 | 2 | 63*a*, 65*b* |
| 30 | Zinc . . | 65·38 | 7 | 64*a*, 65*e*, 66*b*, 67*d*, 68*c*, 69*g*, 70*f* |
| 31 | Gallium . | 69·72 | 2 | 69*a*, 71*b* |

TABLE XIX—*continued.*

| Atomic Number. | Element. | Atomic Weight. | No. of Atomic Species. | Masses of Isotopes. |
|---|---|---|---|---|
| 32 | Germanium . | 72·60 | 8 | 70*c*, 71*g*, 72*b*, 73*d*, 74*a*, 75*e*, 76*f*, 77*h* |
| 33 | Arsenic . . | 74·96 | 1 | 75 |
| 34 | Selenium . | 79·2 | 6 | 74*f*, 76*c*, 77*e*, 78*b*, 80*a*, 82*d* |
| 35 | Bromine . | $79·91_6$ | 2 | 79*a*, 81*b* |
| 36 | Krypton . | 82·9 | 6 | 78*f*, 80*e*, 82*c*, 83*d*, 84*a*, 86*b* |
| 37 | Rubidium . | $85·4_5$ | 2 | 85*a*, 87*b* |
| 38 | Strontium . | $87·6_3$ | 2 | 86*b*, 88*a* |
| 39 | Yttrium . | $88·9_3$ | 1 | 89 |
| 40 | Zirconium . | $91·2_2$ | 3 (4) | 90*a*, 92*c*, 94*b*, (96) |
| 47 | Silver . . | $107·88_0$ | 2 | 107*a*, 109*b* |
| 48 | Cadmium . | $112·4_1$ | 6 | 110*c*, 111*e*, 112*b*, 113*d*, 114*a*, 116*f* |
| 49 | Indium . . | 114·8 | 1 | 115 |
| 50 | Tin . . | $118·7_0$ | 11 | 112, 114, 115, 116*c*, 117*f*, 118*b*, 119*e*, 120*a*, 121*h*, 122*g*, 124*d* |
| 51 | Antimony . | $121·7_6$ | 2 | 121*a*, 123*b* |
| 52 | Tellurium . | 127·5 | 3 | 126*b*, 128*a*, 130*c* |
| 53 | Iodine . . | 126·93 | 1 | 127 |
| 54 | Xenon . . | 130·2 | 9 | 124, 126, 128, 129*a*, 130, 131*c*, 132*b*, 134*d*, 136*e* |
| 55 | Cæsium . | $132·8_1$ | 1 | 133 |
| 56 | Barium . | $137·3_6$ | 1 (2) | (136), 138 |
| 57 | Lanthanum . | $138·9_0$ | 1 | 139 |
| 58 | Cerium . . | 140·13 | 2 | 140*a*, 142*b* |
| 59 | Praseodymium | $140·9_2$ | 1 | 141 |
| 60 | Neodymium . | $144·2_7$ | 3 (4) | 142, 144, (145), 146 |
| 80 | Mercury . | $200·6_1$ | 7 | 196*g*, 198*d*, 199*c* 200*b*, 201*e*, 202*a*, 204*f* |
| 82 | Lead . . | $207·2_1$ | 3 (4) | 206*b*, 207*c*, 208*a*, (209) |
| 83 | Bismuth . | $209·0_0$ | 1 | 209 |

atomic number, tin having the largest number of all, eleven. The numerous instances of so-called *isobares* are also noteworthy, *viz.*, those in which atomic species differing in chemical character have the same atomic weight. A list

TABLE XX

*Radioactive Isotopes (including the Inactive Species of Elements Nos. 81–92)*

| Atomic Number. | Element. | Number of Isotopes. |
|---|---|---|
| 81 | Thallium . . | 4 |
| 82 | Lead . . . | 8 |
| 83 | Bismuth . . | 5 |
| 84 | Polonium . . | 7 |
| 86 | Emanation . . | 3 |
| 88 | Radium . . | 4 |
| 89 | Actinium . . | 2 |
| 90 | Thorium . . | 6 |
| 91 | Protoactinium . | 3 |
| 92 | Uranium . . | 2 |

of the isobaric atomic species hitherto detected in the non-radioactive elements is contained in Table XXI.

TABLE XXI

*Isobaric Atomic Species*

| 40 | 65 | 69 | 70 | 71 | 74 | 75 | 76 | 77 | 78 | 80 | 82 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| A | Cu | Zn | Zn | Ga | Ge | Ge | Ge | Ge | Se | Se | Se |
| Ca | Zn | Ga | Ge | Ge | Se | As | Se | Se | Kr | Kr | Kr |
| 86 | 112 | 114 | 115 | 116 | 121 | 124 | 126 | 128 | 130 | 136 | 142 |
| Kr | Cd | Cd | In | Cd | Sn | Sn | Te | Te | Te | Xe | Ce |
| Sr | Sn | Sn | Sn | Sn | Sb | Xe | Xe | Xe | Xe | Ba | Nd |

Since isotopes are identical in all properties determined by the electrons surrounding the nucleus, it is plain that a *separation of isotopes* can be effected only by such methods as are based on differences in nuclear properties. On the other

hand, the ordinary chemical methods of separation will be of no avail, especially those based on differences in solubility and vapour pressure, in which the principal part is taken by the outer electrons. One of the main possible ways in which isotopes may be separated utilizes the fact that the *velocity of diffusion* and *of evaporation* depend on the atomic mass. Starting with 20,000 litres of hydrochloric acid gas, *Harkins* and *Broeker* [5] have actually succeeded, by diffusion experiments, in effecting a *displacement* of 0·055 units *in the atomic weight of chlorine*, as determined with a quantity of 10 grams of the element. Again, *Brönsted* and *Hevesy*, by the method of successive evaporations, have been able to prepare *two species of mercury* which differ in density by 1 in 2,000, so that their mixture weights differ by 0·1 units.[6] It is, however, too much to hope at present for a preparation of pure isotopes in this manner, for the yield becomes smaller and smaller as the separation proceeds.[7]

The differences in nuclear mass and structure between isotopes also find expression, to a small degree, in *spectral* properties. The conception of the nucleus as a point charge is only approximately correct and the electrical field surrounding the nucleus may exhibit small differences, with varying composition of the nucleus, despite equality of the nuclear charge number. To this is most probably due the fact established by *Merton*,[8] that the wave-length of the spectrum line of *lead* at 4,058 Å. proves to be larger by 0·01 Å. for uranium lead, and smaller by 0·002 Å. for thorium lead, than in the case of common lead. The phenomenon of isotopy finds far stronger expression in molecular spectra than in optical spectra, as we shall later discuss at length.

## § 123. The Packing Effect.

Whereas an explanation for the large deviations of atomic weights from *whole numbers* was afforded by the discovery of isotopy, it has been found that the *small*

[5] W. D. Harkins and C. E. Broeker, *Phys. Rev.*, **15** (1920), p. 74; *Zeitschr. f. Phys.*, **50** (1928), p. 537.

[6] For further details, *vide* Hevesy and Paneth's "*A Manual of Radioactivity*" (Oxford: Clarendon Press, 1926), Ch. xix.

[7] At the present time there does not yet seem to be any possibility of adapting the analytical process of the positive ray method to a method of preparation in which the ions of the isotopes generated in the discharge tube might be separated, in virtue of their different deflections in both an electric and a magnetic field, and then collected.

[8] T. R. Merton, *Proc. Roy. Soc.*, *A* **100** (1921), p. 84.

*deviations*, which exist even with the mass-numbers of the *individual* atomic species, are apparently due to the so-called *packing effect*, in accordance with the Law of the *Mass of Energy*. The *internal energy of atomic nuclei* is so large that the equivalent mass may exert a noticeable influence on the atomic weight. This appears from the fact that, according to § 89 eqn. (2), the kinetic energy of a single $\alpha$-particle, expelled from an atomic nucleus, amounts to $1{\cdot}3 \times 10^{-5}$ ergs, to which corresponds an equivalent mass of $1{\cdot}4 \times 10^{-26}$ grams, or almost 1 % of the mass of a hydrogen atom.

If now the internal energy of a nucleus be $E$, a positive or negative sign being taken according as energy is consumed or liberated in the synthesis of the nucleus, *i.e.*, in packing together its constituents, the *nuclear mass* will appear *increased* (or, in the case of a negative sign, decreased) by an amount $E/c^2$, as compared with the sum of the masses of the component protons and electrons. In consequence of this effect, the atomic weight ($A'$) referred to hydrogen will *differ* from an integral value by the amount

$$\Delta A' = \frac{L\,E}{c^2}, \tag{1}$$

where $L$ denotes the Loschmidt number. Conversely, we can, of course, calculate the internal energy with the help of this equation. For example, the quotient of the atomic weights of *helium* and *hydrogen* is equal to 3·97, so that for helium $\Delta A'$ is equal to $(-0{\cdot}03)$, and we find for the internal energy of the helium nucleus

$$E_{\mathrm{He}} = -4{\cdot}5 \times 10^{-5}\ \text{ergs}. \tag{2}$$

The internal energy of an $\alpha$-particle is therefore about *three times as large* as its kinetic energy for the greatest velocity observed in $\alpha$-rays.

It follows from the negative sign in eqn. (2) that, on *building up a helium nucleus* from its components, *i.e.*, on forming helium from hydrogen, *energy is liberated*. Since there are $1{\cdot}5 \times 10^{23}$ atoms in 1 gram of helium, $6{\cdot}8 \times 10^{18}$ ergs or, approximately, $1{\cdot}6 \times 10^{11}$ calories would be liberated on forming 1 gram. In their great internal energy lies the explanation of the fact that the $\alpha$-particles themselves remain intact, even when they disrupt other atomic nuclei.

The perfecting of mass spectroscopy enabled *Aston*, in 1927, to calculate the internal nuclear energy for a large number of atomic species.[1] As a measure of this energy

[1] F. W. Aston, *Proc. Roy. Soc. A.*, **115** (1927), p. 487.

Aston introduced the so-called *packing ratio*, obtained on *dividing* the deviation of the atomic weight, referred to oxygen, *by the number of protons*. The calculated packing ratios are contained in Table XXII. Hydrogen has the largest value of packing ratio, and the lithium isotopes the next two largest values, although these are only about a

TABLE XXII

*Packing Ratios of Atomic Species (after Aston [2])*

| Atom. | Packing Ratio × 10.[4] | Atom. | Packing Ratio × 10.[4] |
|---|---|---|---|
| H | 77·8 ± 1·5 | $Cl_{37}$ | − 5·0 ± 1·5 |
| He | 5·4 ± 1 | $A_{40}$ | − 7·2 ± 1 |
| $Li_6$ | 20·0 ± 3 | As | − 8·8 ± 1·5 |
| $Li_7$ | 17·0 ± 3 | $Kr_{78}$ | − 9·4 ± 2 |
| $B_{10}$ | 13·5 ± 1·5 | $Br_{79}$ | − 9·0 ± 1·5 |
| $B_{11}$ | 10·0 ± 1·5 | $Kr_{80}$ | − 9·1 ± 2 |
| C | 3·0 ± 1 | $Br_{81}$ | − 8·6 ± 1·5 |
| N | 5·7 ± 2 | $Kr_{82}$ | − 8·8 ± 1·5 |
| O | 0·0 | $Kr_{83}$ | − 8·7 ± 1·5 |
| F | 0·0 ± 1 | $Kr_{84}$ | − 8·2 ± 1·5 |
| $Ne_{20}$ | 0·2 ± 1 | $Kr_{86}$ | − 8·2 ± 1·5 |
| $Ne_{22}$ | (2·2 ?) | I | − 5·3 ± 2 |
| P | −5·6 ± 1·5 | $Sn_{120}$ | − 7·3 ± 2 |
| $Cl_{35}$ | −4·8 ± 1·5 | $Xe_{134}$ | − 5·3 ± 2 |
| $A_{36}$ | −6·6 ± 1·5 | $Hg_{200}$ | + 0·8 ± 2 |

quarter of the hydrogen value. Up to neon the packing ratio is positive,[3] and then from phosphorus onwards it is *negative*, but becomes positive again for mercury. If the packing ratio be represented as a function of the atomic weight, we obtain, as Aston has shown, a fairly regular curve which exhibits a minimum for iron.[4] We find also that, quite generally, the packing effect is larger for elements of odd than for those of even atomic number.

[2] The values given for lithium are due to J. L. Costa, *Ann. der Phys.*, **77** (1925), p. 425.

[3] This means that, for example, in the formation of an oxygen nucleus from four $\alpha$-particles energy would be *liberated*.

[4] Aston refers the wide occurrence of iron to this minimum.

## § 124. Disintegration of the Nucleus, and Transformation of the Elements.

Some few years after the first observation of radioactive radiation, closer investigation led to the discovery of the *transformation of elements*. In 1900 *Rutherford* observed that the radioactive element thorium is constantly evolving a gas that is itself radioactive, and which he called *thorium emanation;* and a similar process was shortly afterwards discovered by *Dorn* to be going on with *radium*. We know now that thorium emanation and radium emanation are isotopes of the rare gas, the nuclear charge number of which is 86; the elementary nature of the emanation was manifest from its *characteristic spectrum*. In 1902 *Rutherford* and *Soddy* developed the so-called *disintegration theory* to explain the formation of the emanations. This theory is based on the assumption that the atoms of radioactive substances undergo a series of successive disintegrations; and that the expulsion of the projected particles in the form of radioactive radiation changes the chemical character of the residual atom.[1]

From the point of view of the nuclear theory of the atom later developed, we must obviously refer the processes of disintegration to the *atomic nucleus* itself; and then, since the *nuclear charge number* determines the *chemical character*, the phenomenon of the *transformation of radioactive elements* at once becomes intelligible. The expulsion of an $\alpha$- or a $\beta$-particle by the disintegrating atom distinguishes between two varieties of transformation, known briefly as the $\alpha$- and the $\beta$-transformation. Since an $\alpha$-particle possesses a positive charge of two elementary quanta and a mass of four hydrogen atoms (§ 89), an *alpha-transformation reduces the atomic number by two*, and the *atomic weight by four*. On the other hand, since $\beta$-particles are identical with electrons, a *beta-transformation raises the atomic number by one, without noticeably altering the atomic weight*. Hence, in the *periodic system* of the elements, an $\alpha$-transformation *displaces* the element affected *through two places to the left*, whilst a $\beta$-transformation displaces it *through one place to the right*. Actually, the emanation resulting, together with $\alpha$-rays, directly from radium is found to have an atomic weight four units

[1] It followed, moreover, that radioactivity is a property of the atoms themselves from the fact that it was found to be quite independent of the kind of chemical compound concerned, and also that it is entirely unaffected by temperature changes of a thousand degrees, whereas the velocity of most chemical reactions is doubled by a rise in temperature of only ten degrees.

lower than that of radium,[2] *viz.*, 222 as compared with 226 ; and, whereas radium belongs to the second group of the periodic system, *i.e.*, the alkaline earth group, the emanation (a rare gas) belongs to the zero group, or, in other words, to the eighth group.

The *formation of helium during the disintegration of the emanation which is itself radioactive* was first demonstrated in 1903 by *Ramsay* and *Soddy*, using a *spectroscopic method.* This was the first instance of the observation of the generation from another element of an element already known to chemists. The actual *identity of the atoms of helium formed* and the *α-particles* expelled during the disintegration of radioactive substances (at least, after neutralization) was proved in 1909 by a beautiful experiment carried out by *Rutherford and Royds.* In this experiment, a substance emitting α-rays was contained in a closed glass tube, the wall of which was not thicker than about one-hundredth of a millimetre, so that the α-rays could pass through it. This tube was placed inside another, thick-walled, glass tube, which was likewise closed, and after some time had elapsed Rutherford and Royds were actually able to demonstrate spectroscopically the presence of helium in the space between the tubes.

On closer investigation, it was found that a preparation containing 1 gram of radium (*e.g.*, in the form of 1·314 grams of $RaCl_2$) generates in a year about 167 mm.$^3$ of helium, corresponding to a mass of 0·0000298 grams (the density of helium being taken as 0·000179).[3] Since the mass of a helium atom is equal to the atomic weight (4·00) divided by the Loschmidt number, we obtain the number of helium atoms generated in a year by 1 gram of radium, on dividing the above-given number (0·0000298) by four, and then multiplying by the Loschmidt number ($6{\cdot}062 \times 10^{23}$).[4] In this way we find that 1 gram of radium produces $4{\cdot}52 \times 10^{18}$ helium atoms in a year.

Again, the accurate counts made by the scintillation method (already mentioned in § 89) have shown that a gram

[2] The density, and hence the atomic weight, of radium emanation were first determined by Gray and Ramsay in 1902, a measured volume of it being weighed by means of a special micro-balance.

[3] The density of monatomic helium is only twice that of diatomic hydrogen.

[4] Since, according to Avogadro's Law (to be discussed in a subsequent chapter), the number of molecules in 1 cm.$^3$ at normal pressure and 0° is the same for all gases, *viz.*, $2{\cdot}708 \times 10^{19}$, the number of monatomic helium molecules contained in 0·167 cm.$^3$ is found simply by multiplying together the two numbers just given.

of radium emits $3{\cdot}72 \times 10^{10}$ $\alpha$-particles in one second.[5] This number is four times as large if the preparation is in equilibrium with the three disintegration products which are formed from radium, and also emit $\alpha$-rays [*cf.* the remarks in § 125 in connection with eqn. (12)]. Since there are 86,400 seconds in a day, and 365 days in a year, the number of $\alpha$-particles ejected by the radium preparation in a year is equal to $3{\cdot}72 \times 10^{10} \times 4 \times 365 \times 86{,}400$, or $4{\cdot}69 \times 10^{18}$. The good agreement between the calculated numbers of the $\alpha$-particles emitted and of the accumulated helium atoms forms an important empirical confirmation of the disintegration theory. Conversely, the observation of the *rate of formation of helium* from radium, together with a *count of the $\alpha$-particles*, enables the *Loschmidt number* to be *determined directly*, although not very accurately.

In every radioactive substance, the ratio of the number of atoms disintegrating in a second to the total number of atoms of the substance present is found to represent a characteristic constant. This ratio is called the *disintegration constant*, on the essential assumption that the *number of atoms disintegrating* in an element of time is always *proportional to the number of atoms available*. Denoting the disintegration constant by $\lambda$ and the number of atoms present by $N$, we have, therefore, for the number of atoms disintegrating in an element of time ($\Delta t$) the relation

$$\Delta N = -\lambda N \Delta t. \tag{1}$$

On replacing the differences by differentials and integrating, we obtain the equation

$$N = N_0 e^{-\lambda t}, \tag{2}$$

where $N_0$ denotes the number of atoms at the time $t = 0$. The time ($T$) in which $N$ diminishes to one-half of $N_0$, half of the atoms originally present having by then disintegrated, is known as the *half-value period*, or period of half change. Its connection with the disintegration constant follows from the relation

$$\frac{1}{2} = e^{-\lambda T};$$

whence

$$T = \frac{\log_e 2}{\lambda} = \frac{0{\cdot}693}{\lambda}. \tag{3}$$

[5] The exact determination of this number is due to *R. W. Lawson* and *V. Hess*, *Sitz. Ber. Akad. Wien, math.-nat. Klasse, Abt. IIa*, **127** (1918), p. 599.

The reciprocal value of the disintegration constant is called the *period of average life ;* it is larger than the half-value period, being obtained on dividing the latter by 0·693.

The disintegration constant of *radium* may at once be calculated from the number of α-particles emitted in a second. For we obtain the number of atoms in 1 gram of radium simply by dividing the Loschmidt number by the atomic weight of radium (226). This gives $2{\cdot}68 \times 10^{21}$; and, since $3{\cdot}72 \times 10^{10}$ α-particles are emitted in one second, while the same number of radium atoms disintegrate, the value resulting for the disintegration constant of radium is

$$\lambda\,(\mathrm{Ra}) = \frac{3{\cdot}72 \times 10^{10}}{2{\cdot}68 \times 10^{21}} = 1{\cdot}39 \times 10^{-11}\ \mathrm{sec.}^{-1}. \tag{4}$$

It follows from eqn. (3) that the *half-value period of radium* is given by

$$T\,(\mathrm{Ra}) = 4{\cdot}99 \times 10^{10}\ \mathrm{sec.} = 1{,}580\ \mathrm{years}. \tag{5}$$

For many radioactive substances, the half-value period (to values as low as about 0·002 sec.) can also be derived from observations on the *diminution in intensity of the radioactive radiation effects*. Thus, denoting the intensity of the radioactive radiation measured by any apparatus [6] at an arbitrary time and at the time $t = 0$ by $I$ and $I_0$ respectively, we have, as is also found experimentally, the relation

$$I = I_0\, e^{-\lambda t}. \tag{6}$$

In the case of radium emanation, the value found in this way for the exponential constant is precisely the same as that given for the disintegration constant by the scintillation method. At the same time, this evidently proves that the intensity of the radioactive radiation at any moment is proportional to the number of atoms of the radioactive substance in existence at that moment.

### § 125. Radioactive Equilibrium.

In the following considerations, we shall apply the fundamental equation of the disintegration theory [§ 124 eqn. (2) ] to the special case in which the *product of the transformation* is itself radioactive, but with a much *smaller life period than the parent substance*. Such, for example, is the case with radium emanation and radium, the half-value periods being in the approximate ratio 1 : 150,000.

[6] The simplest method is to measure the conductivity acquired by air as a consequence of the ionizing effect of the radioactive radiation.

Let $\lambda$ be the disintegration constant of the parent substance, and $\lambda'$ that of the transformation product. We shall suppose $N$ atoms of the former and $N'$ atoms of the latter to be present at a given moment. At an arbitrary time $t = 0$, let the values of $N$ and $N'$ be represented by the constants $N_0$ and $N_0'$.

We consider now an arbitrarily chosen element of time $\Delta t$; during that interval, atoms of the transformation product will *disintegrate* and, *simultaneously*, new atoms of the product will *be generated* by the disintegration of the parent substance. The number of disintegrating atoms is $\lambda' N' \Delta t$; the number of newly generated atoms $\lambda N \Delta t$.

If now, as we assumed, the parent substance is very long-lived as compared with the transformation product, then, even after an interval considerably longer than the half-value period of the transformation product, the quotient $N/N_0$ will not differ noticeably from unity.[1] Hence, when we are considering the state of the transformation product, we may always replace the variable quantity $N$ by the constant $N_0$. The change in the number of the atoms of the transformation product in an element of time ($dt$) is accordingly given by the relation

$$\frac{dN'}{dt} = -\lambda' N' + \lambda N_0 . \tag{1}$$

In order to integrate this equation, it is best to begin by differentiating again with respect to the time: denoting the differential coefficient $dN'/dt$ by $\varphi$, for the sake of brevity, we obtain then

$$\frac{d\varphi}{dt} = -\lambda' \varphi . \tag{2}$$

It follows immediately from this that (replacing $\varphi$ by the derivative $dN'/dt$)

$$\frac{dN'}{dt} = C \cdot e^{-\lambda' t}. \tag{3}$$

By means of eqn. (1), we can easily determine the value of the constant $C$, for, at the time $t = 0$, it must be equal to the differential coefficient $dN'/dt$. In accordance with eqn. (1) we have

$$\left|\frac{dN'}{dt}\right|_{t=0} = -\lambda' N_0' + \lambda N_0 . \tag{4}$$

[1] In a system composed of radium and emanation we might take one year as the time.

The right-hand side of this equation therefore represents the constant $C$, so that eqn. (3) may be written in the form

$$\frac{dN'}{dt} = (-\lambda' N_0' + \lambda N_0)\, e^{-\lambda' t}. \tag{5}$$

Integrating this again, we have

$$N' = N_0'\, e^{-\lambda' t} - \frac{\lambda}{\lambda'} N_0\, e^{-\lambda' t} + C'. \tag{6}$$

The integration constant $C'$ can be determined from the equality of the numbers $N'$ and $N_0'$ at the time $t = 0$. Introducing the time $t = 0$ into eqn. (6), we find

$$N_0' = N_0' - \frac{\lambda}{\lambda'} N_0 + C', \tag{7}$$

whence

$$C' = \frac{\lambda}{\lambda'} N_0. \tag{8}$$

Eqn. (6) may therefore be written in the form

$$N' = N_0'\, e^{-\lambda' t} + \frac{\lambda}{\lambda'} N_0 (1 - e^{-\lambda' t}). \tag{9}$$

In this equation let us give $t$ a value which is large compared with the half-value period of the transformation product, but small compared with the half-value period of the parent substance. For such a value of $t$, the expression $e^{-\lambda' t}$ may simply be put equal to zero. Eqn. (9) then becomes

$$N' \lambda' = N_0 \lambda. \tag{10}$$

If, however, this relation be satisfied, then by eqn. (1)

$$\frac{dN'}{dt} = 0. \tag{11}$$

The number of atoms of the transformation product must therefore be constant because, in a given time, just as many new atoms of the transformation product are generated as disintegrate. The system consisting of the parent substance and its transformation product is then said to be in *radioactive equilibrium*. Even if equilibrium did not exist at the time $t = 0$, it must nevertheless be attained after an interval which is large as compared with the half-value period of the transformation product, provided that, during this interval,

there be no external interference with the system (such as the removal of the transformation product as formed).

We shall now suppose the transformation product in its turn to generate a third active substance, the latter generating a fourth product, and so on. Then, when radioactive equilibrium exists, the following relation must hold between the parent substance and all the transformation products :—

$$N_0 \lambda = N' \lambda' = N'' \lambda'' = N''' \lambda''', \text{ etc.} \tag{12}$$

Since the number of atoms of a radioactive substance disintegrating in unit-time is given by the product of the disintegration constant and the number of atoms present, *equal numbers of atoms of all substances* which are in *radioactive equilibrium* with each other *disintegrate* in the same time. This explains why a radium preparation ejects *four times* as many $\alpha$-particles as radium itself, for it contains, besides radium, three transformation products emitting $\alpha$-rays (Rn, RaA, RaC), in equilibrium with the radium.

Eqn. (12), moreover, permits us to derive the *disintegration constant of the long-lived parent substance* from the known disintegration constant of any transformation product. For example, it has been found that, in all uranium minerals, there is a content of $3{\cdot}4 \times 10^{-7}$ grams of *radium* for every gram of *uranium*. Since the atomic weights of uranium and radium are 238 and 226 respectively,[2] it follows that, in radioactive equilibrium, corresponding to $n$-atoms of uranium we have $n \times 3{\cdot}4 \times 10^{-7} \times 238/226$, or $n \times 3{\cdot}6 \times 10^{-7}$ atoms of radium. We obtain then the half-value period of uranium on dividing that of radium (1,580 years) by $3{\cdot}6 \times 10^{-7}$; the *half-value period of uranium* is thus found to be 4,400 million years. This value agrees well with that obtained directly by the scintillation method. Counts have shown that 1 gram of uranium emits $2{\cdot}3 \times 10^4$ $\alpha$-particles in a second, only half of which, however, arise from uranium I, the other half being generated by uranium II, which, as we shall see in the next section, is always present in radioactive equilibrium with uranium I, though in much smaller amount. Since the atomic weight of uranium is 238, 1 gram contains about $2{\cdot}55 \times 10^{21}$ atoms. The disintegration constant is therefore found on dividing the half of $2{\cdot}3 \times 10^4$ by this number, the result being

$$\lambda\,(\mathrm{U}) = 4{\cdot}5 \times 10^{-18}\ \text{sec.}^{-1}.$$

[2] It is not necessary to take into consideration uranium II, which is always contained as an isotope in the mixed element uranium, for it only forms about 0·025% of the whole element.

It follows from this, according to § 124 eqn. (3), that the half-value period of uranium amounts to $4{\cdot}9 \times 10^9$ years, in good agreement with the value derived indirectly.

The ratio between the half-value periods of *radium* and *radium emanation* is also found to obey eqn. (12) satisfactorily. One gram of radium is in equilibrium with 0·663 mm.$^3$, or $6{\cdot}56 \times 10^{-6}$ grams of emanation, this quantity, which serves as a unit of measure, being known as a "*curie*."[3] The half-value period of emanation must hence be equal to that of radium multiplied by $6{\cdot}56 \times 10^{-6}$, and by the ratio between the atomic weights of radium and emanation, *i.e.*, by 226/222.[4] In accordance with § 124 eqn. (5), we find thus for the half-value period of radium emanation $3{\cdot}33 \times 10^5$ sec., while direct observation on the decrease in the intensity of the radiation gives a value of $3{\cdot}30 \times 10^5$ sec., or 3·82 days.

Finally, we will apply eqn. (9) to the special case in which, at a given moment which we choose as zero on the time scale, the *transformation product is separated from the longer-lived parent substance* with which it is supposed, till then, to have been in equilibrium. From the moment of separation, we have therefore to distinguish between two preparations. In the first, the transformation product which was completely removed at the time $t = 0$ continues to form, whilst in the second preparation, which contains no parent substance, the number of atoms of the transformation product in it decreases continuously, in consequence of their progressive disintegration. Let $N_1'$ and $N_2'$ be the numbers of atoms of the disintegration product in the two preparations at any given time.

It follows from eqn. (9) that, since $N_1'$ is to be made equal to zero at the time $t = 0$, the number $N_1'$ is given by

$$N_1' = \frac{\lambda}{\lambda'} N_0 (1 - e^{-\lambda' t}) . \tag{13}$$

On the other hand, at the time $t = 0$ the number $N_2'$ is equal to $N_0 \lambda/\lambda'$, because of the equilibrium existing up to that

[3] The unit of measure used for radium emanation concentrations is the *Mache* unit; this corresponds to an emanation content of $3{\cdot}64 \times 10^{-10}$ curie/litre. (The quantity of emanation then contained in one litre will produce, on transference to an electroscope, a saturation current of $10^{-3}$ electrostatic units a second.) The strongest therapeutic springs have an emanation content of 2,000 or 3,000 Mache units.

[4] For $n$ radium atoms are in equilibrium with $n \times 6{\cdot}56 \times 10^{-6} \times 226/222$ emanation atoms.

time. Hence, at an arbitrary time [in accordance with § 124 eqn. (2)],

$$N_2' = \frac{\lambda}{\lambda'} N_0 e^{-\lambda' t}. \tag{14}$$

Adding eqns. (13) and (14), we find

$$N_1' + N_2' = \frac{\lambda}{\lambda'} N_0; \tag{15}$$

*i.e.*, the *sum of the atoms of the transformation product in the two preparations* is *not affected by the separation*, but remains *constant*.

An instructive example of the fulfilment of this law is

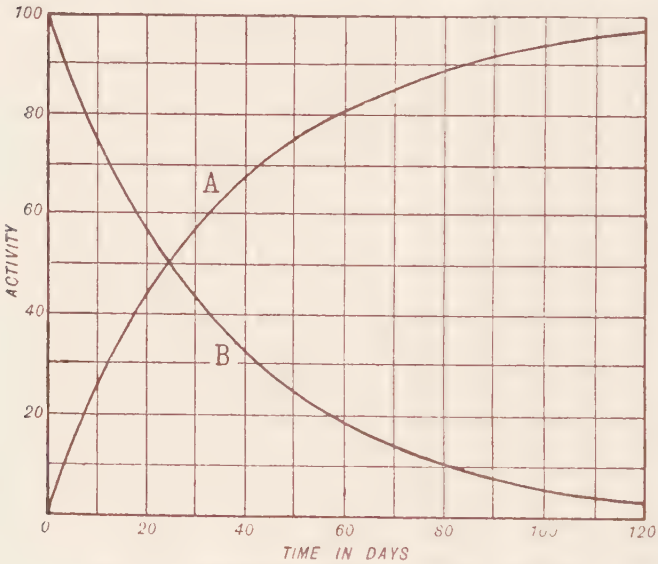



FIG. 100.—Separation of uranium X from uranium.

afforded by the *separation of uranium X from uranium*. Uranium X is of far shorter life than uranium, its half-value period amounting to only 23·8 days.[5] Since uranium emits $\alpha$-rays, but uranium X $\beta$-rays, the intensity of the $\beta$-radiation emitted by the separated preparations forms a measure of the number of uranium X atoms present.[6] In Fig. 100 curve A shows the *increasing* intensity of the $\beta$-radiation

[5] The small admixtures of uranium $X_2$ and uranium II can be disregarded, and we may speak simply of uranium and uranium X. *Cf.* § 126.

[6] The $\alpha$-radiation can be cut out by absorption.

from the "recovering" uranium preparation, and curve B the *decreasing* intensity in the separated uranium X preparation. At every instant the sum of the ordinates of the two curves has the same value, namely, the initial value (supposed equal to one hundred).

### § 126. The Transformation Series.

At the present time, 40 radioactive elements,[1] or, as they are briefly termed, "*radioelements*," with the most various life periods are known in physics; the half-value periods range from $10^{-11}$ sec. to $10^{12}$ years. Apart from potassium and rubidium, from which a weak $\beta$-radiation has been detected, the radioelements are distributed over the last twelve places in the periodic system with the ordinal numbers 81 to 92, omitting, however, the 85th and 87th places. Exact investigations have shown that the radioelements (except potassium and rubidium) can be arranged in *three transformation series*, in such a way that each element in such a series arises from the preceding element through either an $\alpha$- or a $\beta$-transformation. The three series are known as the uranium-radium series, the actinium series, and the thorium series, though the second is obviously only a branch of the first.

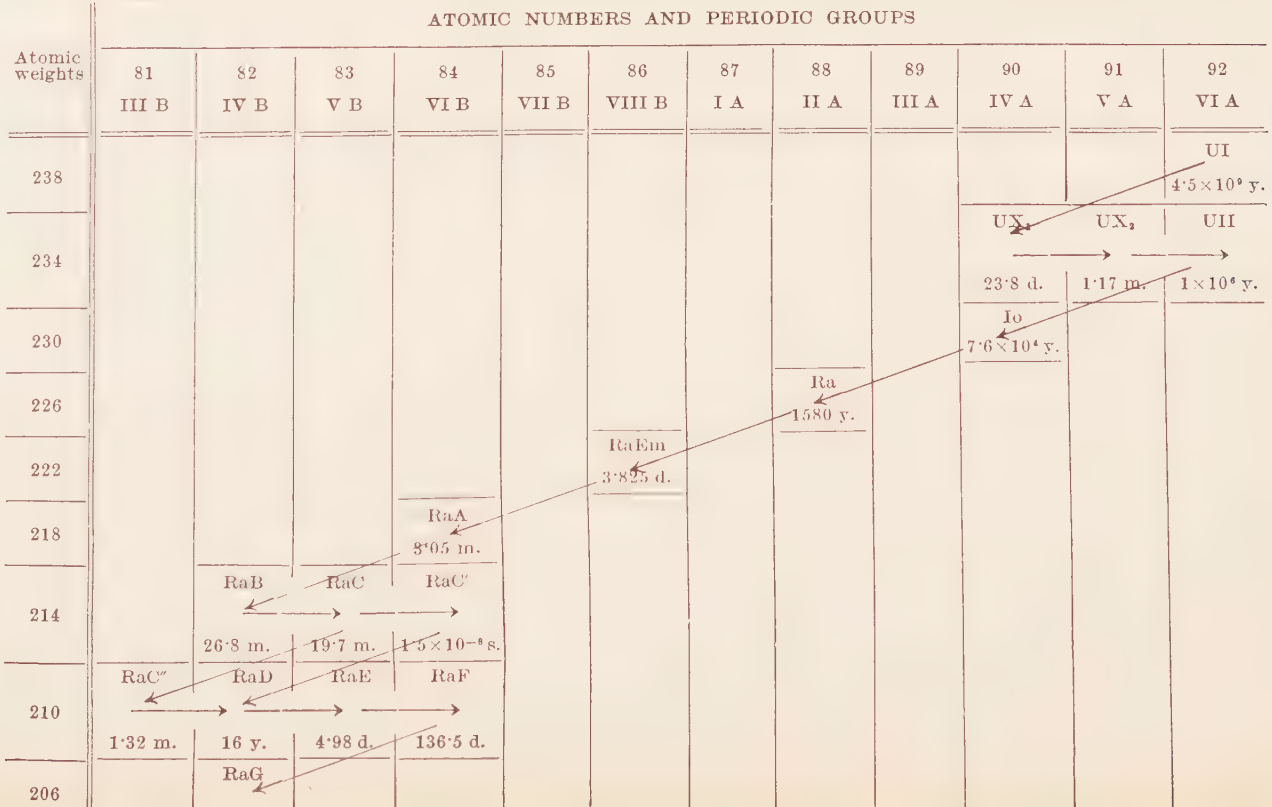
The *uranium-radium series* is represented in Table XXIII, in which elements with the same atomic weight are situated in the same row, and elements with the same atomic number in the same column; the groups of the periodic system corresponding to the atomic numbers are also shown. An $\alpha$-transformation, in which the atomic weight decreases by four, is indicated by an oblique arrow, and a $\beta$-transformation, in which the atomic weight remains unaltered, is shown by a horizontal arrow. The half-value period is given under each element (here, and in the two following tables, years, days, hours, minutes, and seconds are denoted respectively by the letters *y*, *d*, *h*, *m*, and *s*).

As can be seen from Table XXIII, *uranium* is a *mixed element* with *two isotopes*, the heavier of which undergoes one $\alpha$- and two $\beta$-transformations to form the lighter. One gram of uranium I, however, is in equilibrium with only 0·25 mgm. of uranium II, and, indeed, with only $1{\cdot}5 \times 10^{-11}$ grams of uranium $X_1$, and $4{\cdot}9 \times 10^{-16}$ grams of uranium $X_2$. Hence from the point of view of practical chemistry, the mixed element can be regarded as being pure uranium I. Uranium II

[1] Not including the stable end products of the transformations.

## TABLE XXIII

*The Uranium-Radium Series*



cannot be separated from uranium I, with which it is isotopic ; on the other hand, uranium $X_1$ and uranium $X_2$, which are isotopes of thorium and protoactinium, can be separated by suitable precipitation methods.

An $\alpha$-transformation leads from uranium II to the long-lived parent substance of radium, *ionium*. Inasmuch as this is an isotope of thorium, it cannot be prepared in a pure state, but only in the form of an enriched thorium preparation, to the extent of 30%.[2] Radium gives rise to *radium emanation* on an $\alpha$-transformation.[3]

Radium emanation produces the so-called *short-lived active deposit of radium*, consisting of five substances, RaA, RaB, RaC, RaC′ and RaC″, the genetic relations of which can be seen from Table XXIII. These substances are all solid, and are mostly charged positively immediately on their production, so that they are readily deposited on negatively charged plates. The five elements are so short-lived that they can be observed only in mixtures, though the proportions of the mixtures may be varied at will. Analyses of the curves representing the decrease in activity of such mixtures have made it possible to determine the half-value periods of the individual constituents of the short-lived active deposit. The smallest of the half-value periods have been indirectly determined on the basis of the relation between the life period and the range of the emitted $\alpha$-rays, which we shall discuss in the following section.

At radium C, the transformation series *branches*, a $\beta$-transformation leading to the extraordinarily short-lived RaC′, whereas an $\alpha$-transformation leads to RaC″. The $\alpha$-transformation, however, is much rarer than the $\beta$-transformation; of 10,000 RaC atoms, only 4 undergo the transformation into RaC″.

When a sufficiently large quantity of emanation (such as half a cubic millimetre) disintegrates inside a closed glass vessel, a fine deposit which has the properties of a *species of lead* (known as *radium D*) is formed on the walls. The transformation law shows that radium D must result from an $\alpha$-transformation of RaC′ (which emits $\alpha$-rays), and also by a $\beta$-transformation from RaC″. Radium D has the comparatively large half-value period of 16 years ; it changes after emitting a weak $\beta$-radiation into an isotope of bismuth,

[2] Whereas the atomic weight of pure thorium is 232·1, the combining weight of the thorium preparation richest in ionium is 231·5.

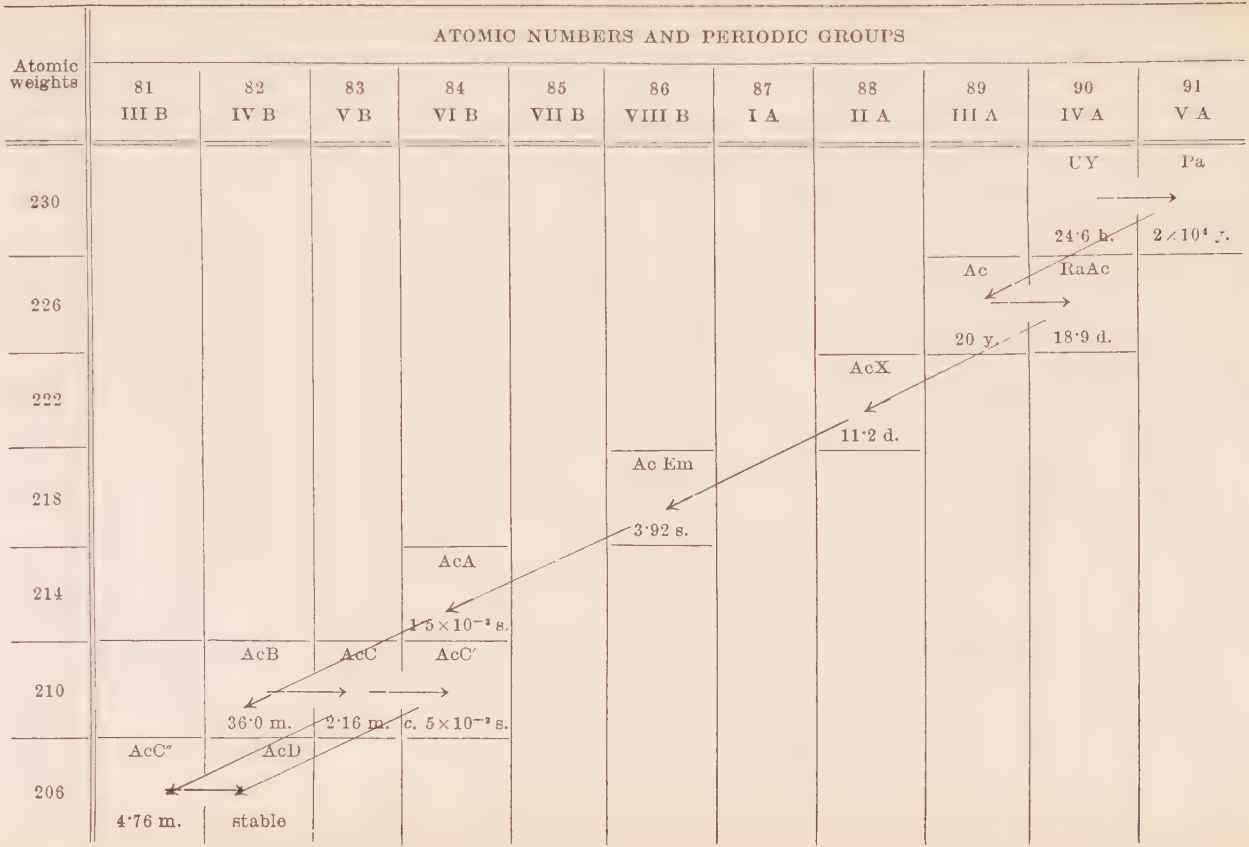
[3] Since this is a rare gas, not chemically combined but only occluded in radium salts, it can easily be liberated on heating.

known as radium E, from which another $\beta$-transformation gives rise to RaF, or *polonium*. Polonium emits $\alpha$-rays, and when it has decayed no new rays can be detected. Hence the element produced from polonium after an $\alpha$-transformation and known as RaG, which (according to the transformation law) is a *species of lead*, the so-called *uranium lead*, must be assumed to be the *end product of the uranium-radium series.*

Quite analogous to the radium series is the *actinium series*, represented in Table XXIV. There is no doubt about the sequence from protoactinium to the presumably final product actinium D, isotopic with lead. *Protoactinium*, which was first prepared in a pure state by *Grosse* [4] in 1927, is a very long-lived element, with a half-value period of 20,000 years ; the disintegration product, arising from an $\alpha$-transformation, is *actinium*, the half-value period of which is much smaller, amounting only to 20 years. One $\beta$- and two $\alpha$-transformations lead from actinium to *actinium emanation*, the life period of which, however, only amounts to a few seconds. Actinium emanation gives rise to the *active deposit* of actinium. The sequence of transformations in this deposit much resembles that of the radium active deposit. Whereas, however, almost the whole of the RaC atoms undergo a $\beta$-transformation, just the converse is the case with the atoms of AcC. Of 10,000 AcC atoms, 9,968 emit $\alpha$-rays and form AcC″, whilst only 32 undergo the $\beta$-transformation into AcC′. AcC′ and AcC″ lead directly to the final stable product, AcD, whereas the corresponding product radium D is weakly radioactive.

Evidence that the *actinium series* probably represents a *branch of the uranium series* is to be found in the observed fact that the analogous members of the actinium series occur in all uranium minerals to the same extent, being responsible for about three *per cent.* of the total activity. This means that the uranium series branches in such a way that 97 out of every 100 atoms follow the transformation into radium, and 3 the transformation into actinium. At what point the branching occurs is, however, not yet certain. On the grounds of the transformation law alone, it is probable that the parent substance of protoactinium is *uranium Y*, an isotope of thorium, emitting $\beta$-rays, which is present in uranium and contributes three *per cent.* towards the activity, though the relative masses are only in the ratio of $2 \times 10^{-14}$ to 1. It is

[4] A. v. Grosse, *Chem. Ber.*, **61** (1928), p. 233; *Naturwiss.* **15** (1927), p. 766.


ATOMIC NUMBERS AND PERIODIC GROUPS
Atomic weights
81 III B
82 IV B
83 V B
84 VI B
85 VII B
86 VIII B
87 I A
88 II A
89 III A
90 IV A
91 V A
230
226
222
218
214
210
206
UY
Pa
24·6 h.
2 × 10⁴ y.
Ac
RaAc
20 y.
18·9 d.
AcX
11·2 d.
Ac Em
3·92 s.
AcA
1·5 × 10⁻³ s.
AcB
AcC
AcC′
36·0 m.
2·16 m.
c. 5 × 10⁻³ s.
AcC″
AcD
4·76 m.
stable

further possible that uranium Y, like ionium, is produced from uranium II by an $\alpha$-transformation, so that a branching involving two $\alpha$-transformations would occur at this radio-element. The atomic weights of the actinium series are still only hypothetical.

A further *branching in the uranium series* gives rise to *uranium Z*, 35 out of 10,000 uranium $X_1$ atoms being transformed into uranium Z atoms, with an emission of $\beta$-rays; a second $\beta$-transformation leads to uranium II. Uranium Z is an isotope of protoactinium, has a half-value period of 6·7 hours, and contributes about 0·35 *per cent.* towards the activity of uranium, in which it is always present, though the mass ratio is only $6 \times 10^{-16}$ to 1.

The probable *independence of the thorium series from the uranium series* is indicated by the fact that both uranium minerals almost free from thorium, and thorium minerals almost free from uranium, are known. The thorium series is shown in Table XXV.

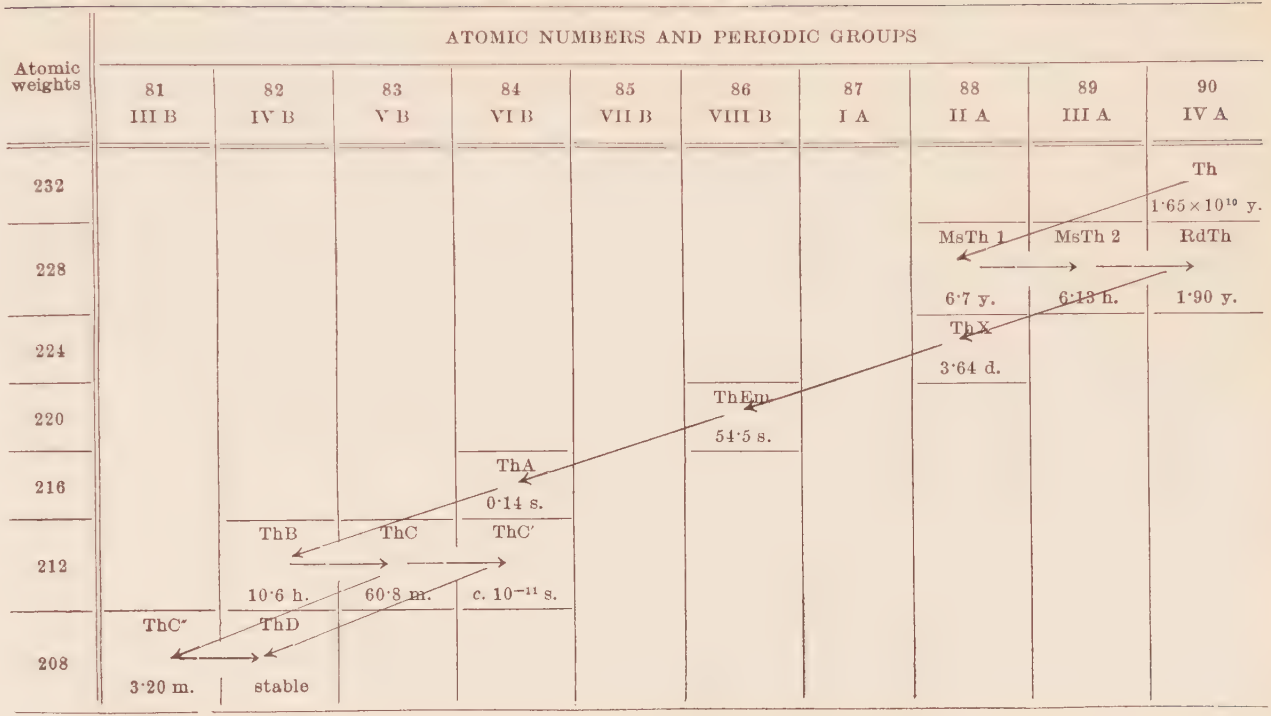
From the initial substance of the thorium series, *thorium*, an $\alpha$-transformation leads to the relatively long-lived *mesothorium* 1. Since its half-value period amounts to 6·7 years, and in all thorium minerals there are $4 \times 10^{-10}$ grams of mesothorium 1 to 1 gram of thorium, it follows from the law of radioactive equilibrium that the half-value period of thorium must be $1{\cdot}65 \times 10^{10}$ years. Pure mesothorium 1 cannot be prepared from minerals, because it is an isotope of radium, and all thorium minerals always contain a little uranium, and thus radium.[5] The pure substance can only be obtained in minute quantities by the disintegration of thorium salts which have been previously purified from uranium and its transformation products.

From mesothorium 1, two successive $\beta$-transformations lead through the short-lived mesothorium 2 to *radiothorium*, which is an isotope of thorium with an atomic weight smaller by four units, and with a half-value period of nearly two years. From radiothorium, two $\alpha$-transformations give rise to the short-lived *thorium emanation*, which in its turn produces the *active deposit of thorium*. Considerable analogy exists between this and the deposits of radium and actinium, the final product here being thorium D, which is

[5] Since 1 gram of thorium is in equilibrium with $4 \times 10^{-10}$ grams of mesothorium, and 1 gram of uranium with $3{\cdot}4 \times 10^{-7}$ grams of radium, the quantities of mesothorium and radium contained in a thorium mineral will be approximately the same when uranium is present in the mineral to the extent of even only one in a thousand of the thorium content.

## TABLE XXV

*The Thorium Series*



isotopic with lead. Approximately 65% of thorium C undergoes a $\beta$-transformation and 35% an $\alpha$-transformation. About one ten-thousandth of the atoms of ThC disintegrate with the emission of $\alpha$-rays of abnormally long range (11·3 cm.), so that at least a three-fold branching of the thorium series must be assumed to occur at ThC.

The *final product* of all three transformation series is represented by different isotopes of *lead*. Consequently, both in uranium and in thorium minerals, appreciable quantities of lead must accumulate in the course of a very long time; and hence a determination of the lead content enables us to fix the *age of the minerals*, and therewith the age of the *geological formations* in which the minerals are found. Since $n \times 1{\cdot}5 \times 10^{-10}$ atoms out of $n$ uranium atoms disintegrate in the course of one year, while the ratio of the atomic weight of lead to that of uranium is 206 : 238, 1 gram of uranium must produce annually $1{\cdot}3 \times 10^{-10}$ grams of uranium lead. We obtain therefore the approximate age (in years) of a mineral on multiplying the amount of uranium lead, corresponding to 1 gram of uranium, by the reciprocal of $1{\cdot}3 \times 10^{-10}$ or, roughly, by 8,000 million.[6] A ten *per cent.* content of uranium lead thus indicates an age of about 800 million years. Determinations of age must, however, be based only on the content of uranium lead (RaG), from which a common contamination of ordinary lead must be distinguished. Inasmuch, however, as the atomic weights of uranium lead and ordinary lead are different (206·0) and 207·2), a determination of the atomic weight of the mixture of RaG and ordinary lead at once enables us to calculate the proportions of the two substances.[7] In this way it has been found that the *age* of uranium minerals ranges from about 300 million to 2,000 million years.[8]

Whether the uranium and thorium series, both of which presumably end with lead, begin with elements of still

[6] In an accurate calculation it must be remembered that the uranium has also decreased in amount in the course of time; allowing for this, the ages found for the minerals are somewhat smaller.

[7] For example, whereas the Morogoro ore found in East Africa gives an atomic weight of 206·05, and can thus only contain uranium lead, it follows from the atomic weight of 206·4, found for the lead in uraninite from North Carolina, that 70% of the lead has been formed from uranium, and 30% is due to an impurity.

[8] Thorium minerals are frequently of secondary origin, so that the ages found for them cannot always be ascribed to the geological formations in which they occur. The ages calculated for thorium minerals range from 10 million to 600 million years.

higher atomic weight than uranium and thorium is an open question. But it is unlikely that such higher elements can be found, for, if they were shorter-lived than uranium or thorium, then, in accordance with what has just been said about the age of minerals, they would have disintegrated long ago, seeing that a formation of new elements after the solidification of the earth is scarcely conceivable. On the other hand, if their life periods were sufficiently long, they must occur in recognizable quantities in uranium and thorium minerals, as a consequence of radioactive equilibrium, in which case they could hardly have escaped detection up till now. It does not, however, appear impossible for elements with higher atomic weights than that of uranium to exist, if they belong to transformation series other than the three described, which may have hitherto eluded discovery simply because of the rarity of their members.

Members of such transformation series, as yet unknown, are possibly represented by those isotopes of *potassium* and *rubidium* to which is due the $\beta$-radiation discovered in these two metals by *Campbell* in 1907. It is true that, both with potassium and with rubidium, it is still an open question whether one of the two known isotopes is radioactive, or whether the activity is caused by a third unknown isotope. If the former be the case, it may be estimated from the hardness of the $\beta$-rays emitted that the half-value periods of rubidium and potassium range from about $10^{11}$ to $10^{12}$ years.[9]

It was not until after the first observations of radioactivity in 1896 that by far the greater number of the radioelements were discovered; before that date only uranium and thorium were known, apart from potassium and rubidium. Of the radioelements discovered after that time, five represent newly discovered elements in the sense that they are the principal representatives of places in the periodic system which had previously been empty. These five are *radium* and *polonium*, discovered in 1898 by *M.* and *Mme. Curie*; *actinium*, discovered by *Debierne* and *Giesel* in 1899; *emanation*, discovered in 1900 by *Rutherford* and *Dorn*; and finally, *protoactinium*, discovered by *Hahn* and *Meitner* and by *Soddy* in 1918.[10] The genetic relations between the radioelements became evident on the establishment (in 1913) by

[9] *Cf.* the researches by *A. Holmes* and *R. W. Lawson*, *Phil. Mag.*, (7) 2 (1926), p. 1218.

[10] Before protoactinium, its short-lived isotope uranium $X_2$ (formerly called "brevium") had already been discovered.

*Soddy* and by *Fajans* of the law of $\alpha$- and $\beta$-transformations —even before the discovery of the natural sequence of the elements.

### § 127. Material Radiation of Radioelements.

The most striking phenomenon due to the disintegration of the nucleus, which, indeed, first drew the attention of physicists to that disintegration, is the *radioactive radiation* emitted by the disintegrating atoms, part of which is *material* in nature and part of the nature of *electromagnetic waves*. The radioactive rays arise either as so-called *primary rays from the atomic nucleus*, or as *secondary rays* in the *electron groups* surrounding the atomic nucleus.

Foremost among the primary rays are the $\alpha$- and $\beta$-particles ejected by the disintegrating atomic nucleus. The *initial velocities of the $\alpha$-particles* represent *constants* characteristic of the radioelements emitting them. Expressed in cm./sec., the initial velocities lie between $1{\cdot}40 \times 10^9$ for uranium I, and $2{\cdot}06 \times 10^9$ for thorium C′ ; the abnormally rapid $\alpha$-rays, which are emitted by about a ten-thousandth part of the disintegrating thorium C atoms, have a velocity of $2{\cdot}3 \times 10^9$.

As was mentioned in an earlier section (§ 89), the *range* of the $\alpha$-rays in any gas is *proportional to the cube of the initial velocity ;* the same relation also exists between the distance yet to be covered at any moment and the constantly decreasing velocity at that moment. The ranges in air at atmospheric pressure and 0° likewise represent constants characteristic of the radioelements which emit $\alpha$-rays ; they vary from 2·53 cm. for uranium I to 8·17 cm. for thorium C′. The exceptionally rapid $\alpha$-rays emitted by a ten-thousandth part of the disintegrating thorium C atoms have a range of 11·3 cm. Furthermore, the range varies directly as the absolute temperature of the gas, and inversely as its pressure, so that the range of $\alpha$-rays *in vacuo* is found to be extremely large. Referred to RaC′, the range in air at 0° and atmospheric pressure is connected with the initial velocity by the relation

$$v^3 = 1{\cdot}076 \times 10^{27} R_0 , \tag{1}$$

or,

$$v = 1{\cdot}025 \times 10^9 R_0^{1/3} . \tag{2}$$

Ranges may also be determined directly by a photo-

graphic method based on Wilson's "fog track" device, previously described. Fig. 101 is a reproduction of such a photograph, by Lise Meitner, of the effect produced by the radioactive deposit of thorium. With a suitable arrangement of screens, all the paths can be photographed in the same plane. In the figure we can distinguish $\alpha$-rays both of 4·8 cm. and of 8·6 cm. range, while two isolated paths due to rays of 9·5 and 11·3 cm. range can be seen. If air be replaced by another gas, it is true that four other values are obtained for the range, but the same ratio exists between these as in the case of air. By taking several hundreds of similar photographs it was, moreover, ascertained that one ray of 11·3 cm. range occurs to about every 5,000 rays of 8·6 cm. range, whereas the corresponding ratio for rays of 9·5 cm. range is only about 1 in 15,000. The same proportions are also observed in gases other than air.

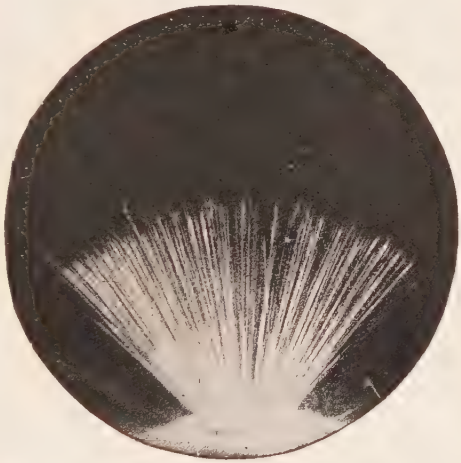


Fig. 101.—$\alpha$-rays from the radioactive deposit of thorium (after Meitner, "*Atomvorgänge und ihre Sichtbarmachung*": Stuttgart: Enke, 1926).

A simple relation connects the ranges of $\alpha$-radiators with their *life-periods*. As was found by *Geiger* and *Nuttall* in 1911, the *logarithms of the disintegration constants* are *linear functions of the logarithms of the ranges*, within each transformation series. We have the relation

$$\log \lambda = A + B \log R_0 , \tag{3}$$

where $B$ is a universal constant, but $A$ has a definite value for each of the three transformation series (*cf.* Fig. 102). With logarithms to the base 10, $B$ is equal to 60 and, for the uranium-radium series, $A$ is equal to $(-41{\cdot}5)$. By means of eqn. (3) it has been possible to determine the half-value periods of the exceedingly short-lived radioelements. That eqn. (3) is not generally valid, however, can be seen from the

position of AcX in Fig. 102, the range corresponding to the thorium rather than to the actinium series.[1]

Whereas the ranges are taken as characteristic constants for the radioelements emitting $\alpha$-rays, the *absorption coefficients in aluminium* are generally taken for those elements which emit $\beta$-rays. Many radioelements possess *several*

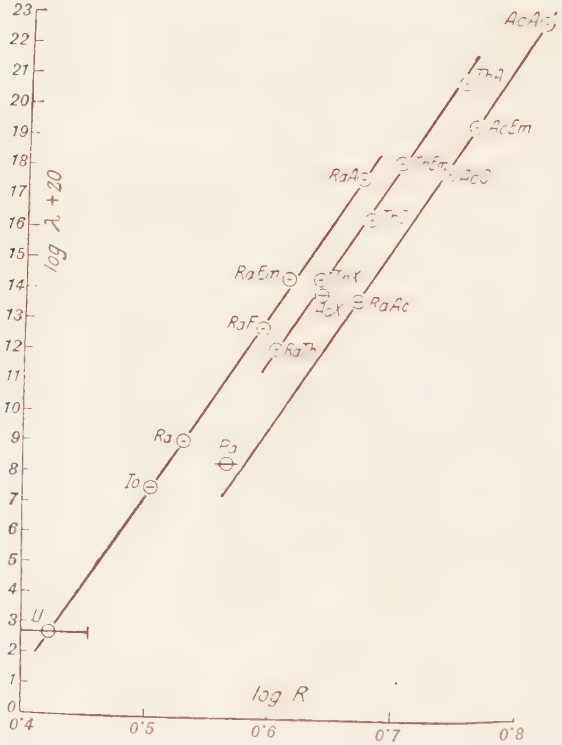



FIG. 102.—Relation between range and disintegration constant. (*Zeitschr. f. Phys.*, vol. 8. Springer, Berlin.)

*different values for the absorption coefficient*, which shows that the $\beta$-radiation emitted by them is composed of rays of different velocities. This is still more clearly demonstrated by the *magnetic spectra* (investigated principally by *Kovarik*), which are based on the different deflections of rays of different

[1] As Fig. 102 shows, Geiger and Nuttall's equation does not hold in general for the thorium and the actinium series nearly so accurately as for the uranium-radium series.

velocities in a magnetic field. Thus, if the $\beta$-radiation as deflected by a magnet be allowed to fall on a photographic plate (which is thereby blackened), with a suitable arrangement of the apparatus sharp black lines will be formed on the plate (Fig. 103). Each of these lines corresponds to a definite initial velocity, which can be easily calculated,[2] but some of them are due to the secondary $\beta$-rays (which we shall discuss later), only those of higher velocity being of primary origin. The initial velocities of the $\beta$-rays lie between $9{\cdot}9 \times 10^9$ and $2{\cdot}994 \times 10^{10}$ cm. sec., amounting to between 33 and 99·8% of the velocity of light. The fastest $\beta$-rays have been found in the magnetic spectrum of radium C. In general, large values of the initial velocity are associated with small half-value periods.

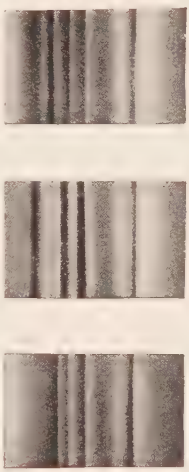


Fig. 103. — Magnetic spectra of $\beta$-rays.

By measuring the charge imparted by $\beta$-rays to an electroscope, the *number of $\beta$-particles emitted in a second* by a $\beta$-radiator can be calculated. It is found that this number is always *equal to the number of atoms undergoing a $\beta$-transformation* in a second. For example, it has been shown that the quantities of the $\beta$-radiating elements RaB and RaC in equilibrium with a gram of radium emit together $7 \times 10^{10}$ $\beta$-particles in a second. Half of this number, however, is nearly identical with the number of $\alpha$-particles expelled by a gram of radium (apart from transformation products), which, as already mentioned, amounts to $3{\cdot}72 \times 10^{10}$. (The agreement is not perfect, because the primary $\beta$-particles are accompanied by secondary particles.)

Like $\alpha$ rays, $\beta$-rays also produce a considerable *ionization* of the air. A $\beta$-particle gives rise to about 10,000 pairs of ions, an $\alpha$-particle to between 100,000 and 250,000 according to its range. Whereas, as previously mentioned (*cf.* the earlier Figs. 61 and 62), $\alpha$-particles generate *linear* streaks in supersaturated vapour, the tracks formed by $\beta$-particles are quite *irregular*, because the much lighter $\beta$-particle (as opposed to the heavier $\alpha$-particle) is thrown out of its path by every collision with an air atom. Furthermore, since the number of air ions produced by a $\beta$ particle is much

[2] *Cf.* § 88 eqn. (1).

smaller, the *path of a β-particle*, photographed with the same magnification, appears *resolved* into individual drops of mist (Fig. 104).

Besides primary β-rays, *secondary* β-rays are also emitted by radioelements. These are the result of a *photo-electric effect* produced in the groups of electrons surrounding the nucleus by the primary γ-rays emitted from the latter. They are, of course, to be observed not only in such radioelements as undergo a β-transformation, but also in those which experience an α-transformation: in point of fact, a weak β-radiation of secondary origin can be detected in the α-radiating elements, radium, radioactinium, and radio-



FIG. 104.—Path of a β-particle (after Meitner, "*Atomvorgänge und ihre Sichtbarmachung*": Stuttgart: Enke, 1926).

thorium. *Secondary γ-rays*, due to primary β- or primary α-rays, are simply the *characteristic X-rays* of the elements concerned, which are also to be found in their X-ray spectra.

A part of the material radiation emitted by radioactive substances is formed by the so-called *recoil rays*. In accordance with the Principle of the Conservation of Momentum, every atom which expels an α- or a β-particle must experience a recoil, with a velocity the ratio of which to that of the expelled α- or β-particle is the same as that of the particle's mass to the mass of the remainder of the atom. For example, if an atom of radium A with an atomic weight of 218 expels an α-particle of mass 4 with a velocity of $1{\cdot}69 \times 10^9$ cm. sec., the atom of radium B resulting from the transformation must recoil with a velocity equal to that of the α particle just given multiplied by 4/214. This

velocity is about $3 \times 10^7$ cm./sec., and it is large enough to render the recoil rays recognizable.

The *heat* which is continuously evolved by an active preparation enclosed in a vessel originates in the kinetic energy of the $\alpha$- and $\beta$-rays and the recoil rays in their collision with the walls of the vessel, and also in the absorbed electromagnetic energy of the $\gamma$-rays. It may be calculated from the characteristic constants of the rays, and from the disintegration constants of the radium series, that the heat evolved in an hour by a gram of radium, together with its short-lived transformation products, must amount to 137 calories. Of this quantity of heat 90·1% is due to the $\alpha$-rays, 3·4% to the $\beta$-rays, 4·7% to the $\gamma$-rays, and 1·8% to the recoil rays. In good agreement with the calculated result, the hourly development of heat is found by experiment to be 136 calories.[3] The total quantity of heat which would be evolved by a gram of radium on complete disintegration amounts to $3{\cdot}7 \times 10^9$ calories. It is, perhaps, of interest to note that this quantity is about a million times as large as that developed on the transformation of a gram of "electrolytic gas" into water, although this chemical reaction is marked by an unusually large evolution of heat.

### § 128. Gamma-Radiation and Nuclear Levels.

In contradistinction to $\alpha$- and $\beta$-radiations, which are composed of material particles, the $\gamma$-radiation emitted by the radioelements represents, as has frequently been mentioned, a *radiation of light-quanta* of especially *high frequency*. In fact, $\gamma$-rays are simply *extremely hard X-rays*. It is possible to measure wave-lengths down to about 30 X (0·03 Å) by means of *interference phenomena*, as has been done by *Rutherford*, *Andrade*, *Kovarik*, *Thibaud*, and others. Still smaller wave-lengths have been attained by indirect methods, one of which, devised by *Ellis* [1] and by *Lise Meitner*,[2] is particularly noteworthy: this method utilizes *magnetic $\beta$-spectra*.

The wave-lengths of the $\gamma$-rays of a radioelement can be derived from its magnetic spectrum, provided that the *energy-levels* of an element *isotopic* with it are known from

[3] The continuous evolution of heat has also been detected in the cases of polonium, thorium, and uranium. A gram of uranium in equilibrium with its disintegration products develops about a ten-thousandth of a calorie in an hour.

[1] C. D. Ellis, *Proc. Roy. Soc.*, *A* **99** (1921), p. 261.

[2] L. Meitner, *Zeitschr. f. Phys.*, **9** (1922), pp. 131 and 145.

X-ray spectroscopic measurements. According to the fundamental relation obtained in an earlier section (§ 90), the kinetic energy of a *secondary* $\beta$-particle is given by the equation

$$L = h\nu - W, \tag{1}$$

where $\nu$ is the frequency of the primary $\gamma$-rays and $W$ the work necessary to remove the $\beta$-particle from its usual orbit "to infinity."

We shall now assume that the magnetic $\beta$-ray spectrum of a radioelement contains two lines corresponding to velocities $v'$ and $v''$. The difference in these values is obviously explained by the fact that the secondary $\beta$-particle is *liberated from a different level* in either case. If, however, the liberation is effected in each instance by $\gamma$-rays of the *same frequency*, we find, on applying eqn. (1) to each case and subtracting,

$$L' - L'' = W'' - W'; \tag{2}$$

in calculating the kinetic energy, the dependency of the mass on the velocity must be taken into account [according to § 88 eqn. (6)].

Now if both secondary $\beta$-rays really were generated by primary $\gamma$-rays of the same frequency, we see that the energy difference $(W'' - W')$ divided by $h$ gives the frequency of a known *X-ray line* of the isotopic element. This being the case, however, we can immediately find out from what levels the secondary $\beta$-particle was liberated in the two instances, provided we know the energy-levels of the isotopic element. We have only to replace $W$ in eqn. (1) by the spectral term in question multiplied by $h\,c$, when, since the kinetic energy of the secondary $\beta$-particle is known, the frequency of the primary $\gamma$-ray can at once be calculated from eqn. (1).

Determinations of the wave-lengths of $\gamma$-rays lead to considerably smaller values than have been found for the hardest X-rays. It is, however, impossible that $\gamma$-rays of shorter wave-length than the $K$-radiation of the heaviest elements should be generated by processes in the planetary systems of the atoms; they must instead be regarded as a *nuclear radiation*. The shortest $\gamma$-ray wave-length yet measured amounts to 6 X, or about one-twentieth of the smallest X-ray wave-length [3]; the corresponding potential is about 2,000,000 volts.

[3] *Cf.* the article by L. Meitner on the relation between $\beta$- and $\gamma$-radiation in "*Ergebnisse der exakten Naturwissenschaften*" (Berlin, Springer), **3** (1924), p. 161.

In 1921 *Ellis*[4] made the important discovery that the $\gamma$-lines are governed by very similar *combinational relations* to those holding for optical and X-ray lines. He obtained the following values for six hard $\gamma$-rays of radium B (the corresponding energy values in volts $\times 10^5$ are also shown) :—

(3)

| | | | | | | |
|---|---|---|---|---|---|---|
| $\lambda$ (in X) | 51·9 | 48·8 | 42·3 | 35·4 | 33·9 | 30·8 |
| $E$ (in $10^5$ volts) | 2·385 | 2·529 | 2·918 | 3·492 | 3·639 | 4·000 |

Now the values of the energy can be arranged in *three pairs with constant differences, viz.*,

$$4{\cdot}000 - 2{\cdot}918 = 1{\cdot}082\ ;\ 3{\cdot}639 - 2{\cdot}529 = 1{\cdot}110\ ;$$
$$3{\cdot}492 - 2{\cdot}385 = 1{\cdot}107.$$

From this possibility Ellis derived the scheme of energy-

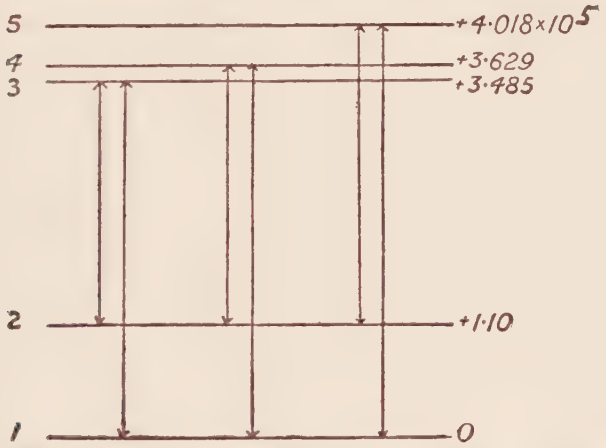



Fig. 105.—Scheme of energy-levels in the radium B nucleus (after Ellis).

levels and corresponding transitions for the radium B nucleus reproduced in Fig. 105. Denoting the levels in order of magnitude of the energy values by 1, 2, 3, 4, 5, he found that the lines previously mentioned correspond to transitions from the third, fourth, and fifth to the first and second levels. Next he calculated the wave-lengths which would correspond to the other transitions, and, as Table XXVI shows, the values so obtained agree well with those actually measured by Rutherford and Andrade.

[4] C. D. Ellis, *Proc. Roy. Soc.*, *A* **101** (1922), p. 1 ; C. D. Ellis and H. W. B. Skinner, *ibid.*, **105** (1925), pp. 60, 165, 185.

Besides the nuclear wave-lengths, other wave-lengths, which *coincide* with those of *X-rays* as derived from other sources, occur in the $\gamma$-radiation. These correspond to *secondary $\gamma$-rays* generated in the planetary system, the shell of the atom concerned, by primary processes in the nucleus. *Lise Meitner* [5] was able to show that, on the $\beta$-disintegration

TABLE XXVI

*Comparison of the $\gamma$-Wave-lengths Calculated from the Level Scheme of the Radium B Nucleus with the Observed Wave-lengths*

| Level combination. | Energy in $10^5$ volts. | $\lambda$ (in $X$) calcd. | $\lambda$ (in $X$) measured. |
|---|---|---|---|
| 5 — 4 | 0·389 | 318 | 324 |
| 5 — 3 | 0·533 | 231 | 229 |
| 4 — 3 | 0·144 | 857 | 853 |
| 2 — 1 | 1·10 | 112 | 115 |

of an atom of nuclear charge number $N$, X-ray lines characteristic of the element $(N + 1)$ are produced, whereas, on an $\alpha$-disintegration, the lines are characteristic of the element $(N - 2)$. This appears to prove that the emission of a secondary $\gamma$-light-quantum is preceded by the *expulsion of a component of the nucleus.*

When $\gamma$-rays impinge on matter of any kind, *secondary electrons* are liberated [6]; since each light-quantum thereby gives rise to one electron, *Hess* and *Lawson* [7] were able to *count* the $\gamma$-light-quanta emitted by a radioactive preparation. *Kovarik* [8] perfected this method by substituting *absolute* for the original relative counts of Hess and Lawson, and he was thus able to show that the *number of $\gamma$-light-quanta emitted* is always *identical with the number of atoms disintegrated.* On arranging two similar counting apparatuses in different directions, but at the same distance, from a source of $\gamma$-rays, he found that the two apparatuses *never*

[5] *Vide* L. Meitner's article, "*Kernstruktur*" in the *Handbuch der Physik* (Berlin; Springer, 1926), Vol. XXIII.

[6] This is, of course, an instance of the well-known photo-effect.

[7] V. F. Hess and R. W. Lawson, *Sitz.-Ber. der Wiener Akad.*, Abt. IIa, **125** (1916), pp. 285, 585.

[8] A. F. Kovarik, *Phys. Rev.*, **6** (1915), p. 426; **8** (1916), p. 574; **13** (1919), p. 272; **14** (1919), p. 179; **23** (1924), pp. 228, 559.

registered *simultaneously*. This remarkable observation of Kovarik's may be regarded as a direct experimental proof of the *quantum* character of $\gamma$-rays.

### § 129. Cosmic Rays.

Although $\gamma$-rays are equivalent to potentials of as much as 2,000,000 volts, yet their wave-lengths are not the shortest known to physicists. Some octaves still higher lie rays which presumably reach the earth from space, and are therefore known as *cosmic rays*.[1] The starting point for their discovery was formed by the observation that, during the ascent of a balloon, the *electrical conductivity of the air*, as measured in a closed vessel, *increases rapidly at a height of some kilometres*. This phenomenon was first noticed by *Hess* in 1912, and by *Kolhörster* somewhat later. Hess also discovered that the conductivity is just as strong by night as by day, so that a radiation emitted by the sun can scarcely be the cause. Since the ordinary radioactive substances cannot be the origin of the conductivity (for then the effect would be stronger on the earth's surface than at a height above it), Hess assumed that an explanation of the remarkable phenomenon must be sought in a new, hitherto unknown, radiation.

An investigation into the nature of Hess's radiation was first accomplished by *Millikan* and *Cameron*,[2] who made the surprising discovery that, in high mountain lakes, these rays are effective to a depth of 70 metres. Since, with respect to absorption, the atmosphere above the lakes was equivalent to a layer of water 7 metres in thickness, the rays can therefore penetrate some 80 metres of water. The equivalent thickness of lead would be some 10 metres, whereas the hardest X-rays can penetrate only to a depth of rather more than 1 cm. in lead. More accurate experiments finally led Millikan and Cameron in 1928 to the important conclusion that cosmic radiation must also be *discontinuous*, so that it must represent a species of line spectrum.[3] Millikan observed that, after penetrating to a certain depth in water, the *absorption coefficient suddenly changes*, and he concluded therefore that the radiation is composed of *constituents with different powers of penetration*. He deduced from his measurements that the cosmic radiation

[1] They are also termed ultra-gamma-rays—a designation which is preferred by those who regard a cosmic origin of the rays as only hypothetical.

[2] R. A. Millikan and G. H. Cameron, *Phys. Rev.*, **28** (1926), p. 85.

[3] *Ibid.*, **31** (1928), p. 921.

consists principally of three bands, the *absorption coefficients* of which, referred to water and 1 metre,[4] are 0·35, 0·08, and 0·04. Since a relation connecting absorption coefficient and wave-length has been tentatively propounded, the corresponding wave-lengths could be calculated. The band with the strongest power of penetration is associated with the *shortest* of all known *wave-lengths*, *viz.*, 0·057 X, and hence, in accordance with § 93 eqn. (3), with an equivalent potential of 216 million volts.

Now the energy liberated on the *synthesis of a helium nucleus* from four protons and two electrons (*cf.* § 123) represents a light-quantum of 0·43 X wave-length : according to Millikan, the corresponding absorption coefficient in water is 0·30 per metre. From this coincidence of values, Millikan concludes that the band of greatest wave-length in the cosmic radiation, which includes about 80% of the rays, is to be referred to the formation of helium. Cosmic rays of larger wave-length, and therefore with a lesser power of penetration, occur only with a smaller intensity ; whilst rays of shorter wave-length are first observed at a *distance of three octaves from* 0·43 *X*. Millikan and Cameron, therefore, assume that the cosmic rays belonging to the bands of shorter wave-length are generated by a *direct, immediate, formation of higher atomic nuclei from protons and electrons.* From the agreement between the calculated and the observed absorption coefficients, Millikan and Cameron think that it may be justifiably concluded that it is a question of the formation of the most common elements, *viz.*, *oxygen*, *silicon, magnesium* and, possibly, *iron.*

It is also of great interest that, according to Millikan, the total energy reaching the earth in the form of cosmic radiation accounts for about one-tenth of the energy coming to the earth as starlight.

### § 130. The Disruption of the Elements.

In radioactive phenomena, the disintegration of the nucleus represents a spontaneously occurring process which cannot be influenced in any way whatsoever. It was, therefore, a discovery of the greatest importance when, in 1919, *Rutherford* for the first time demonstrated an *artificial disruption of an element*, and that an element with a low atomic number.

Rutherford's discovery was based on observations on the

[4] If the absorption coefficient be denoted by $\mu$, the intensity of the radiation, after penetrating through $x$ metres, drops to the amount $I = I_0 e^{-\mu x}$.

origin of the H-rays which we have already mentioned in an earlier section (§ 89). Shortly after the detection of these rays, *Marsden*, in conjunction with *Lantsberry*, was surprised to find that a piece of nickel foil, coated with radium C, produced *scintillations* through *ordinary air* at a distance much greater than the range in air of the $\alpha$-rays emitted from the nickel foil. For, whereas the range of these $\alpha$-rays (actually coming from the disintegration product RaC′) amounts only to 7 cm., scintillations could be observed even at a distance of 40 cm. This was all the more unexpected, because even H-rays generated in hydrogen have a range in air of only 29 cm.

On following up this, at first mysterious, phenomenon, Rutherford made the important discovery that the number of scintillations depended on the substance filling the space between the nickel foil and the screen. When this space was evacuated or filled with carbon dioxide or oxygen instead of with air, there were no scintillations; on the other hand, the number increased on replacing the air by pure nitrogen. At the same time, measurements of the electric and magnetic deflections proved that the long-range rays actually were H-rays, and thus consisted of hydrogen nuclei.

These facts led Rutherford to the important conclusion that the *$\alpha$-particles effect a disruption of the nitrogen nuclei, hydrogen nuclei being expelled* from the disrupted nuclei.

It is found from calculations on scintillation counts that, on an average, only one single hydrogen nucleus is liberated for some hundred thousands of $\alpha$-particles emitted. Since the amount of RaC′ in radioactive equilibrium with a gram of radium ejects $3{\cdot}72 \times 10^{10}$ $\alpha$-particles in a second, a preparation of 1 gram of radium (of which, according to Rutherford, only the RaC′ is of account for the disruption of the nitrogen) will thus liberate about $10^{12}$ hydrogen nuclei annually. Hence, the hydrogen atoms arising on neutralization would, in a year, make up a volume of about a thousandth of a cubic millimetre.

In 1925 *Blackett*,[1] by applying Wilson's method, succeeded in *photographing* also the *disruption of nitrogen nuclei*. Figs. 106 and 107 (Plate IV) show first of all the passage of $\alpha$-particles through air and hydrogen (after Blackett). Furthermore, they illustrate, on the one hand, the *deflection* of the $\alpha$-particle on its *collision* with an atomic nucleus, and, on the other hand, the *recoil* experienced by the nucleus involved; the lighter the nucleus, the larger is the recoil.

[1] P. M. S. Blackett, *Proc. Roy. Soc.*, *A*. **107** (1925), p. 349.

(Hence the recoil tracks in Fig. 107 are larger than in Fig. 106 : the short branches always represent the path of the $\alpha$-particle.) The actual disruption is evident in Fig. 108 (Plate IV). On account of the rarity of the process, photographs of the paths of some 400,000 $\alpha$-particles had to be taken in order to include eight instances of an actual disruption in the exposure : two of these can be seen in Fig. 108. The long thin branch represents the far-reaching track of the *proton* expelled from the disrupted nucleus, and the shorter branch the track of the recoiling *nitrogen nucleus*, deprived of the proton. The absence of any other branch denotes that the $\alpha$-particle, after liberating the proton from the nitrogen nucleus, must coalesce with the latter. Since, in this process, the original nitrogen nucleus exchanges an $\alpha$-particle for a proton, the resulting product, after neutralisation, must be an *isotope of oxygen* with an atomic weight of 17. Such an atom is not otherwise known, and is evidently not stable.

*Rutherford* and *Chadwick* [2] have succeeded, with the $\alpha$-particles of RaC′, in liberating hydrogen nuclei from the elements 5 B, 7 N, 9 F, 10 Ne, 11 Na, 12 Mg, 13 Al, 14 Si, 15 P, 16 S, 17 Cl, 18 A, 19 K. With a large number of other elements (in particular, with C and O), they obtained *negative* results. These were, however, contradicted by *Kirsch* and *Pettersson* who, using other experimental methods found that, with the exception of helium itself, all the light elements and at least ten of the heavier elements (including, for example, Cu and Fe) could be disrupted.[3]

The contradictory results obtained with *carbon* and *oxygen* are of especial importance. It is, indeed, a fundamental question for nuclear physics whether or not it is possible to disrupt atoms of even ordinal number and a *mass number divisible by four*, *i.e.*, atoms which may apparently be constructed *entirely of $\alpha$-particles*. Since isotopy has not been detected in carbon and oxygen, it would seem that such atoms must be susceptible to disruption, provided that the effect can really be obtained with these two elements.[4]

Considerable differences of opinion with respect to some other questions still prevail between the Cambridge school

[2] E. Rutherford and J. Chadwick, *Phil. Mag.*, **48** (1924), p. 509 ; J. Chadwick, *Phil. Mag.* (7), **2** (1926), p. 1056.

[3] H. Pettersson and G. Kirsch, "*Atomzertrümmerung*" (Leipzig ; Akad. Verlagges., 1925).

[4] In the case of sulphur, which can also be disrupted according to Rutherford, either of the two isotopes 33 or 34 may be affected, whereas the most frequently occurring atoms of mass 32 might not be susceptible.



FIG. 106.—Passage of α-particles through air (after Blackett).



FIG. 107.—Passage of α-particles through hydrogen (after Blackett).

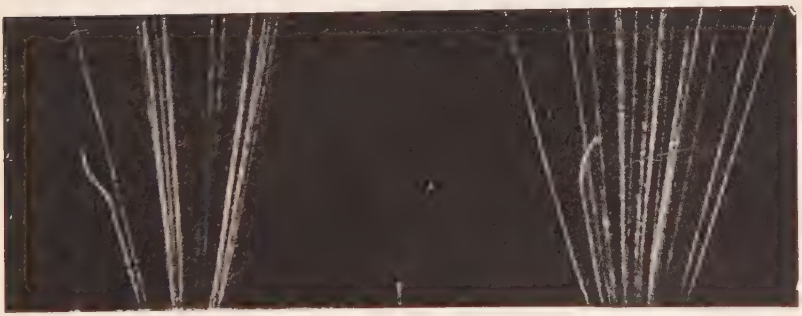


FIG. 108.—Disruption of nitrogen nuclei by α-particles (after Blackett).

[*To face p.* 214.

(Rutherford and his collaborators) and the Vienna school (Kirsch, Pettersson, and their collaborators). The two schools hold quite different views as to the yield of atomic fragments. Whereas Rutherford considers that $\alpha$-rays can exert a disruptive effect only when their range exceeds 4·9 cm., Kirsch and Pettersson are of opinion that, if such a minimum range exists at all, it must be less than a centimetre. Furthermore, it has been found at Cambridge that the atomic fragments themselves must exhibit a minimum residual range, whereas the contradictory observation has been made at Vienna that the number of atomic fragments increases steadily as the range decreases. At the moment it does not seem possible to decide between the opposing points of view, but, at the same time, the divergent results appear to be connected with differences in the experimental methods.[5] Nevertheless, at least the actual phenomenon of atomic disruption may be regarded as conclusively proved with many elements, and hence it represents a certain experimental proof for the fundamental assumption of atomic theory, *viz.*, that, apart from electrons, the atomic nuclei of hydrogen form the sole constituents of all the elements.[6]

## § 131. The Structure of Atomic Nuclei and Harkins' Rules.

Since atomic nuclei are composed of protons and electrons, we may justifiably hope that theoretical physics will succeed in explaining the *structure of atomic nuclei* according to principles similar to those governing the building up of the planetary systems; in other words, we may look forward to the establishment of a *quantum theory of atomic species* as a parallel to the quantum theory of the periodic system. The non-fulfilment of this hope as yet (1929), and the great difficulties standing in the way of a theory of nuclear structure, are apparently due above all to the fact that, as contrasted with the planetary systems which consist only of electrons, atomic nuclei contain *three kinds of constituents*. These are protons, electrons, and $\alpha$-particles (which presum-

[5] The essential difference lies in the adoption by Pettersson and his collaborators of the so-called retrograde method, in which observation is made of secondary particles emitted at an angle of about 150° to the primary $\alpha$-radiation. *Cf.* the summary by G. Stetter, *Physikal. Zeitschr.*, **28** (1927), p. 712.

[6] As long ago as 1815, *Prout* suggested the hypothesis that the atoms of all chemical elements are built up of hydrogen atoms. This concept, which has at times given rise to much thought, lapsed into obscurity when the deviations of the atomic weights from integral values were discovered.

ably possess a certain independence of existence).[1] Although a theory of nuclear structure is still lacking, it has yet been possible to establish a series of important rules affecting this structure, and for these rules we are chiefly indebted to *Harkins* (1915 *ff*).[2]

The first of Harkins' rules states that, in the sequence of the elements, an element of *even ordinal number* is always of *more frequent occurrence* than its two *neighbours with odd*

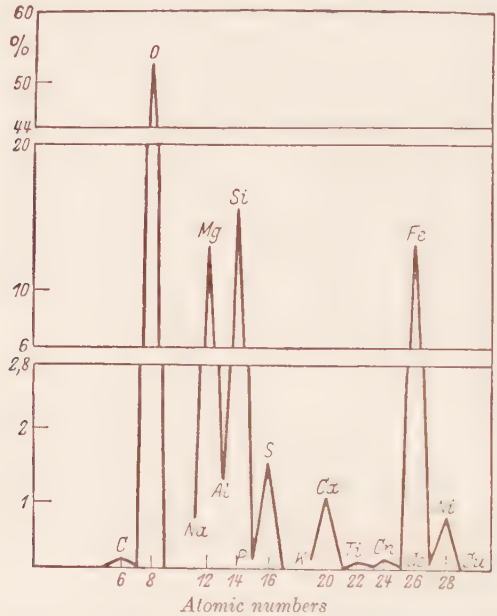



FIG. 109.—Relation between atomic percentages in stone meteorites and atomic numbers (after Harkins, *Zeitschr. f. Phys.*, Vol. 50 : Springer, Berlin).

*ordinal numbers*. This rule appears very clearly in the diagram of Fig. 109, which shows the *atomic percentages* [3] of the elements 5 to 29 in *stone meteorites*. It is noteworthy that Harkins' rule is obeyed even by the *rare earths*, which agree so closely in chemical behaviour that they can be separated only with difficulty. Fig. 110 reproduces the

[1] Another difficulty is that, within the nucleus, the deviations from Coulomb's Law are probably much larger than in the planetary systems.

[2] W. D. Harkins, *Zeitschr. f. Phys.*, **50** (1928), p. 97. Harkins' rules were first published in the *Proc. Nat. Acad., U.S.A.*, **1** (1915), p. 276.

[3] Atomic percentages are used in preference to weight percentages, as giving the commonness of the atomic species directly.

atomic percentages of the different rare earths as given by Harkins for a certain mineral.

The *five most common elements, viz., oxygen, silicon, magnesium, iron*, and *sulphur*, have all *even atomic numbers*, and it is certainly no chance that both the two as yet unknown elements, Nos. 85 and 87, as well as the three most recently discovered, Nos. 43, 61, and 75, have odd atomic numbers. Apart from their more frequent occurrence, the even-numbered elements, as previously mentioned, possess a *larger number of isotopes* than those with odd numbers.

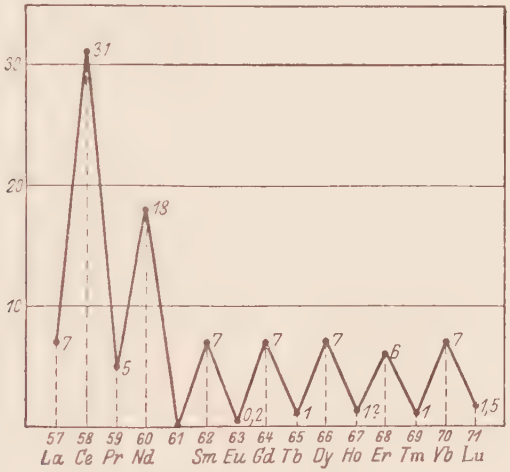



FIG. 110.—Atomic percentages of the rare earths in a mineral (after Harkins, *l.c.*).

Another rule formulated by Harkins states that the *number of nuclear electrons*[4] is *even for a preponderating number of atomic species*. The few atomic species for which this rule does not hold are, moreover, comparatively rare, so that, according to Harkins, $97\frac{1}{2}\%$ of the atomic nuclei in the earth's crust, and $97\%$ of those in stone meteorites, possess an even number of nuclear electrons. Fig. 111 represents the relation between the atomic percentages in stone meteorites and the number of nuclear electrons, again as given by Harkins (the subscripts to the chemical symbols

[4] As already explained, by nuclear electrons we mean the electrons contained in the nucleus as contrasted with the planetary electrons in the atom; the number of nuclear electrons is equal to the difference between atomic weight and atomic number.

indicate the isotope figure—to be defined in a following paragraph).

As regards the *number of protons*, it has been known for some time that elements with an *atomic weight*, and hence with a number of protons, *divisible by four*, are distinguished by their *frequent occurrence*. According to Harkins, the atomic weights of about 90% of all the atoms in stone meteorites, and of 85% of the atoms in the earth's crust, are such as are divisible by four.

With *twelve atomic species in all*, among which helium is the lowest and the calcium isotope of mass 40 the highest, the atomic weight is exactly twice as large as the atomic number; it is larger for all other atomic species, with the sole exception of hydrogen. The difference between the number of nuclear electrons and the atomic number is called the *isotope figure* by Harkins.[5] It is zero for the twelve atomic species mentioned above, and otherwise (except for hydrogen) *always positive*, increasing in general with the atomic number. For uranium the isotope figure amounts to 54 (the atomic weight being 238, the atomic number 92, and hence the number of nuclear electrons 146). As Harkins demonstrated, the *isotope figure* exhibits a *periodicity of four* with respect both to the commonness and also to the number of atomic species, and a secondary periodicity of *two*. It

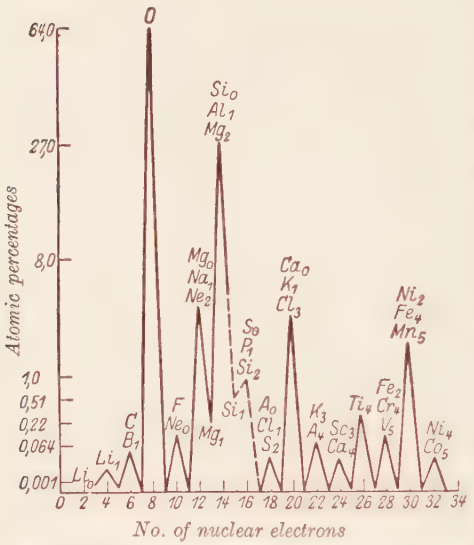



FIG. 111.—Relation between atomic percentages in stone meteorites and number of nuclear electrons (after Harkins, *l.c.*).

[5] If $P$ be the number of protons, $N$ the number of nuclear electrons, and $Z$ the atomic number, we have for the isotope figure $J$ the relation

$$J = N - Z = 2N - P = P - 2Z\ ;$$

for

$$Z = P - N.$$

should also be noted that the *inactive* atomic species can be arranged in *series* which are quite analogous to the *radioactive transformation* series.[6]

Subsequently to Harkins, *Beck*[7] has derived similar rules with the help of a table of isotopes constructed by himself. In particular, Beck has pointed out that the development of the nuclei proceeds in such a way that, after the addition of a certain number of $\alpha$-particles, *two electrons* are *suddenly* absorbed as *binding electrons* (in Beck's terminology) into the nuclear combination. Conversely, the phenomena of radioactive disintegration show that, when the orbit of a binding electron becomes unstable, not only one, but, almost simultaneously,[8] two $\beta$-particles are expelled. The addition of binding electrons to the nuclear combination, and their expulsion therefrom, both *in pairs*, is regarded by Beck as indicating that the phenomenon of *electron spin*, conjointly with *Pauli's principle*, may afford an interpretation for the rules, not only of atomic, but also of nuclear, structure.

[6] *Vide* Harkins, *l.c.*

[7] G. Beck, *Zeitschr. f. Phys.*, **47** (1928), p. 407 ; **50** (1928), p. 548.

[8] In those cases in which a radioelement, itself the product of an $\alpha$-transformation, disintegrates with the emission of $\beta$-rays, the new transformation product is always extremely *short-lived* and again a $\beta$-radiating element. *Cf.* the earlier Tables XXIII–XXV.

## Chapter XVI

# THE MOLECULE

### § 132. Rotation-Vibration Spectra.

Although the *spectra of molecules*[1] are far more complicated than those of atoms, yet it has been found possible to interpret the essential rules of molecular spectra on the supposition that *molecular energy* consists of *three components*, each of which can assume only *discrete* values with *definite differences*. We take as first component the *rotational energy* of the molecule, as second component the *energy of the nuclear vibrations*, and as third component the *energy of the electrons*, due to their orbital motions. As might be expected, the "steps" in the rotational energy prove to be small in comparison with those in the vibrational energy, and the latter small as compared with the steps in the electron energy.

Let us first consider the *rotational energy* ($E_r$). According to an elementary relation of mechanics [2],

$$E_r = \frac{U^2}{2I}, \tag{1}$$

where $U$ denotes the angular momentum and $I$ the moment of inertia about the axis of rotation (which, in the case of a diatomic molecule, is to be regarded as being at right angles to the line joining the atomic nuclei). According to the definition of moment of inertia, we have for a diatomic molecule

$$I = m_1 r_1^2 + m_2 r_2^2, \tag{2}$$

where $r_1$ and $r_2$ denote the distances of the two atoms (of masses $m_1$ and $m_2$) from the centre of gravity. Hence

$$I = \frac{m_1 m_2}{m_1 + m_2} r^2, \tag{3}$$

where $r$ is the distance between the two atomic nuclei.[3]

[1] Monatomic molecules are naturally excluded from these considerations.

[2] *Vide* Vol. I., § 28 eqn. (14) and § 29 eqn. (1).

[3] For $$r_1 = \frac{m_2}{m_1 + m_2} r, \text{ etc.}$$

In the sense of the new quantum mechanics, we have for the angular momentum, according to § 97 eqn. (25),

$$U = \frac{h}{2\pi}\sqrt{l\,(l+1)}\,, \tag{4}$$

where $l$ is a quantum number. Hence

$$E_r = \frac{h^2}{8\,\pi^2\,I}\,l\,(l+1)\,. \tag{5}$$

Since the quantum number $l$ determines the angular momentum, we shall at first consider only those transitions in which $l$ changes by $\pm 1$ [in accordance with § 100 eqn. (7) ]. If we make the simplifying assumption that, on an alteration in the rotational energy, the *moment of inertia* itself remains *unchanged*, we find for the frequency corresponding to a *quantum jump in the rotational energy*

$$\nu_r = \pm\frac{h}{8\,\pi^2\,I}\,[(l+1)\,(l+2) - l\,(l+1)]\,,$$

or,

$$\nu_r = \pm\frac{h}{4\,\pi^2\,I}\,(l+1)\,. \tag{6}$$

A simple special case of changes in molecular energy is formed by those in which the *electron energy* in itself remains *constant*. The manifold of lines corresponding to such changes represents the *rotation-vibration spectrum* of the molecule in question. This spectrum is composed of individual *bands*, each of which is grouped about a definite value of a nuclear vibration frequency ($\nu_s$) according to the formula

$$\nu = \nu_s \pm \frac{h\,z}{4\,\pi^2\,I}\;(z = 1, 2, 3, \textit{etc.})\,. \tag{7}$$

The individual bands must therefore represent *series of equidistant lines*,[4] the central or *zero line* of which will be *missing* ; for, since $z$ is equal to $(l+1)$, the value $z = 0$, and hence $\nu = \nu_s$ is not possible.[5]

Series of *absorption lines* not far in the infra-red (at wave-lengths of a few $\mu$), corresponding to this scheme, have

[4] The equidistance appears as only approximate in Figs. 112 and 113, because the intervals between the lines depend not directly on the frequency differences, but on differences between angles of deflection.

[5] Whereas the earlier quantum theory could only explain the absence of the zero-line with the help of artificial special hypotheses, the necessity for these disappears in the newer theory.

actually been detected and investigated in many substances, especially in the hydrogen compounds of the halogens (HCl, HF, HBr). In Fig. 112 is reproduced a photograph by *Imes* [6] of an *absorption band* belonging to hydrochloric acid and lying between 3 and 4 $\mu$ (30,000 and 40,000 Å.). The ordinates give the absorption expressed as a percentage,

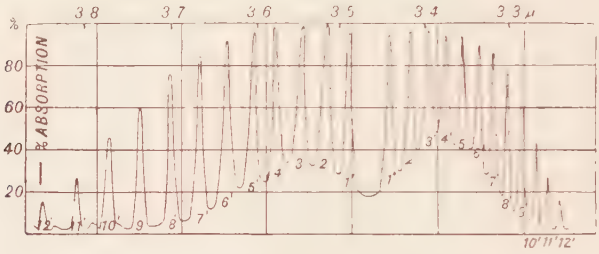



Fig. 112.—Absorption bands in HCl (after Imes).

and the abscissæ angles of deflection measured in connection with the photograph, which was taken with a reflection grating. The maxima along the curve represent the absorption lines, and the corresponding wave-lengths (in $\mu$) are given above them : the figure clearly shows the absence of the zero line.

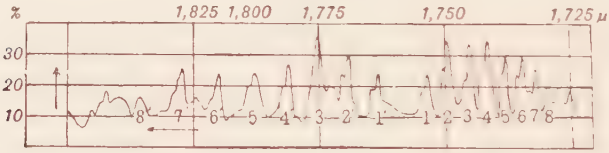



Fig. 113.—Isotope effect in the rotation-vibration spectrum of HCl (after Imes).

The value of $\nu_s$ evidently depends on the *mass* of the oscillating atomic nucleus, and the occurrence of an *isotope effect* may thus be expected in rotation-vibration spectra. *Doublets* of this kind, due to the isotopes of *chlorine*, are in fact clearly recognizable in Fig. 113, in which is reproduced an absorption band of HCl at 1·76 $\mu$, according to Imes.[7] In the neighbourhood of the principal maxima due to the

[6] E. S. Imes, *Astrophys. Journ.*, **50** (1919), p. 251.

[7] This band corresponds to a "harmonic," for the fundamental frequency in Fig. 113 is double that in Fig. 112.

$Cl_{35}$ atoms, the curve exhibits weaker subordinate maxima of smaller frequency, occasioned by the $Cl_{37}$ atoms.[8]

We may, of course, also expect quantum transitions in which the change in the molecular energy is limited to the rotational energy. In such cases we have to deal with a so-called *pure rotation spectrum* which, according to eqn. (7), is represented by the formula

$$(8) \qquad \nu = \frac{z\,h}{4\,\pi^2\,I}\ (z = 1, 2, 3, \textit{etc.})\,.$$

Spectra of this kind have actually been detected and investigated far in the infra-red, notably with *water vapour*.[9]

### § 133. Band Spectra.

Just as the pure rotation spectrum can be superposed upon a nuclear vibration frequency, so also can the whole rotation-vibration spectrum be superposed upon an optical frequency due to the jump of an electron. This process produces a so-called *band system* which is associated with the particular electronic transition. The manifold of all band systems connected with the transitions of electrons in a molecule represents the *band spectrum* of the molecule in question.

In the optical zone, it is true, the rotation-vibration spectrum appears *repeated* only in *distorted forms*. This distortion is due to the *variation* of the distance between the nuclei, and hence of the *moment of inertia of the molecule*, consequent on changes in the electronic configuration. If we denote by $\nu^*$ the frequency associated with the electron transition, by $I$ and $I'$ the moment of inertia before and after, and by $m$ and $m'$ the rotational quantum number before and after the transition, we obtain for the frequencies of the band spectrum the formula

$$(1) \quad \nu = \nu^* + \nu_s + \frac{h}{8\,\pi^2}\left[\frac{m\,(m+1)}{I} - \frac{m'\,(m'+1)}{I'}\right].$$

[8] The difference in frequency between the two lines of an isotope doublet of Cl is theoretically given by

$$\frac{\delta\nu}{\nu} = \frac{1}{2}\left(\frac{1}{35} - \frac{1}{37}\right).$$

It follows that the wave-length difference of the doublet in the band reproduced in Fig. 113 is 13 Å.

[9] *Cf.* the article on infra-red researches by Gerda Laski in "*Ergebnisse der exakten Naturwissenschaften*" (Berlin; Springer, 1924), Vol. III.

To the three transitions

$$m' = \begin{cases} m - 1 \\ m \\ m + 1 \end{cases} \tag{2}$$

correspond within a band *three* so-called *band branches*, which are known as the *positive*, the *zero*, and the *negative branch*; the zero branch is absent in many bands. As we can see from eqn. (1), each of the three branches must be representable in the form

$$\nu = A + B\,m + C\,m^2, \tag{3}$$

where $A$, $B$, $C$ are constants which have the same value for all the lines in a band branch.

If we replace $m$ $(m + 1)$ by $m^2$ in eqn. (1), and likewise $m'$ $(m' + 1)$ by $m'^2$, we pass thereby from the newer, more accurate, quantum mechanics to the older, less exact, quantum theory which, on the basis of empirical rules established by *Deslandres* in 1887, was developed principally by *Heurlinger*, *Kratzer*, and *Lenz* (1919). It is evident that, with this less accurate representation, the constants $A$ and $C$ would have the same values for all three branches of a band, whilst the constant $B$ would have equal but opposite values for the positive and negative branches, and would vanish for the zero branch.

As was first recognized by *Fortrat*,[1] eqn. (3) can be utilized for a simple *graphical representation* of bands. We obtain a *Fortrat diagram* (Fig. 114) on plotting values of $\nu$ as abscissæ, and values of $m$ as ordinates. According to eqn. (3), the values of $\nu$ corresponding to the *integral ordinates* will obviously lie on a *parabola* in the case of each band branch. By drawing perpendiculars from the successive points on these parabolas, we form a picture of the band; and it appears from this that the lines in the band rapidly approach one another towards the projection of the vertex of the parabola. This projection represents the *band head*, which therefore in no way coincides with the *zero position* $(m = 0)$ on the parabola. The zero position is marked on the one hand by the absence of a line (*cf.* the previous section), and on the other hand by the particular intensities of lines with small $m$ values.

It is now of great importance to discover, within a band system, the relations between the different values of the constant $A$ [of eqn. (3)] which determine the zero positions.

[1] R. Fortrat, *Thesis* (Paris; 1914), p. 109.

In this respect it is found that a relation, generally known as the *law of band heads*, is accurately obeyed. This law, which is in principle due to *Deslandres* (about 1900), holds exactly for the *zero positions*, and with great approximation

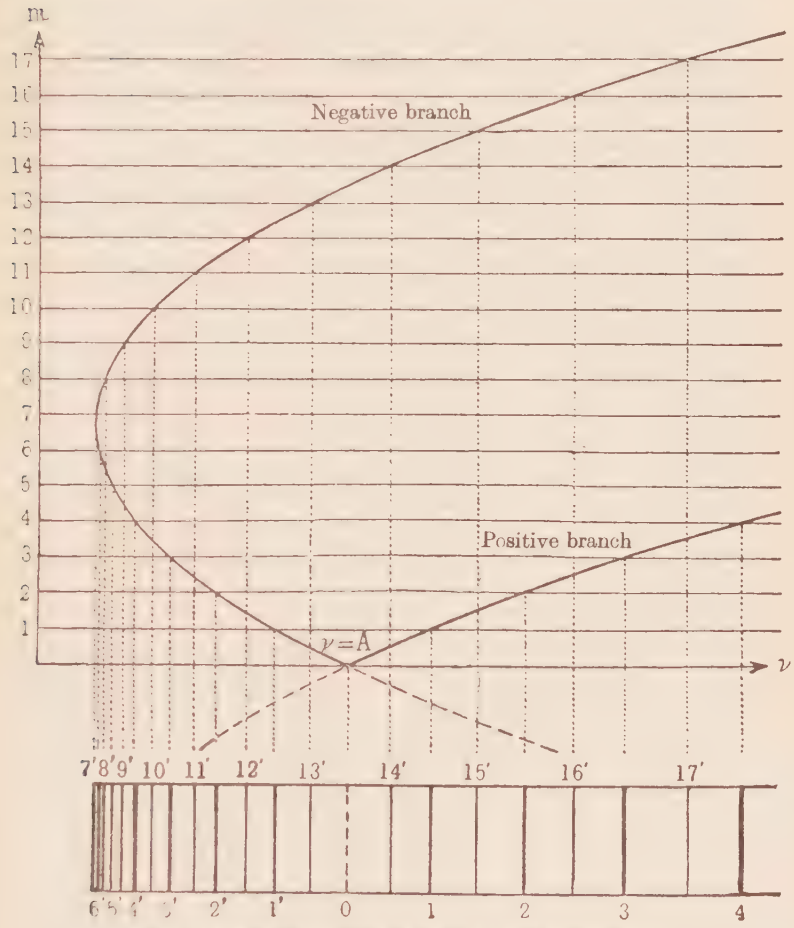



FIG. 114.—Fortrat diagram of a band (after Sommerfeld).

also for the band heads, to which it was first applied. According to it, the possible values of the quantity $A$ within a band system may be represented in the form

$$A = A_0 + (a\,n - b\,n^2) - (a'\,n' - b'\,n'^2), \tag{4}$$

where the four quantities $a$, $b$, $a'$, $b'$ are constants characteristic of the band system.

The constants $A$, which determine the *frequencies of the nuclear vibrations*, can therefore be reduced to a scheme of terms, just as can the spectral frequencies of an atom. In the present case, however, we have to distinguish between *two* different series of terms, corresponding respectively to the *initial* and to the *final state* in the electron transition with which the band system is associated. For the rest, eqn. (4) follows directly from the assumption that the energy of the nuclear vibration can be developed in powers of a quantum number, higher powers than the second being neglected. On the basis of the band head law, the possible values of $A$ can be arranged in a *quadratic scheme*, in which rows correspond to successive values of $n$, and columns to successive values of $n'$ (or *vice versâ*). As is evident from eqn. (4), the differences between the terms in two adjacent rows (or columns) must then be the same. An example of a band head scheme is reproduced in Table XXVII, which

TABLE XXVII

*Example of a Band Head Scheme (Nitrogen Molecule)*

| $n$ \ $n'$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| 0 | 25,797·8 (2,042·4) | 23,755·4 (2,015·9) | 21,739·5 | — |
| | (2,123·5) | (2,123·6) | (2,123·5) | — |
| 1 | 27,921·3 (2,042·3) | 25,879·0 (2,016·0) | 23,863·0 (1,989·6) | 21,873·4 |
| | — | (2,083·7) | (2,082·5) | (2,083·1) |
| 2 | — | 27,962·7 (2,017·2) | 25,945·5 (1,989·0) | 23,956·5 |

represents the scheme of the so-called *cyanogen bands*, actually due to the *nitrogen molecule*. The constancy of the differences between rows and columns (shown in parentheses) is obvious from the table: the values given are all of wave-numbers (frequency divided by the velocity of light). According to eqn. (4), we obtain the following formula for the wave-numbers of the cyanogen bands considered:—

$$(5)\ A = 25{,}797{\cdot}83 + (2{,}143{\cdot}88\,n - 20{\cdot}25\,n^2) - (2{,}055{\cdot}64\,n' - 13{\cdot}25\,n'^2).$$

How accurately this formula is obeyed will appear from Table XXVIII, which is taken from *Kratzer*[2]; the wave-

[2] A. Kratzer, *Physikal. Zeitschr.*, **22** (1921), p. 552; *Ann. der Phys.*, **67** (1922), p. 127.

TABLE XXVIII

*Fulfilment of the Band Head Law by Nitrogen Bands (after Kratzer)*

| $n$ \ $n'$ | 0 | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|---|
| | (3,884) | (4,216) | (4,606) | — | — | — | — |
| 0 | 25,797·83 | 23,755·44 | 21,739·54 | — | — | — | — |
| | 25,797·83 | 23,755·44 | 21,739·55 | — | — | — | — |
| | (3,590) | (3,872) | (4,197) | (4,578) | — | — | — |
| 1 | 27,921·3 | 25,879·0 | 23,863·0 | 21,873·4 | — | — | — |
| | 27,921·38 | 25,878·99 | 23,863·10 | 21,873·71 | — | — | — |
| | — | (3,586) | (3,862) | (4,181) | (4,553) | — | — |
| 2 | — | 27,962·7 | 25,945·5 | 23,956·5 | — | — | — |
| | — | 27,962·04 | 25,946·15 | 23,956·76 | 21,993·87 | — | — |
| | — | — | (3,584) | (3,855) | (4,168) | (4,532) | — |
| 3 | — | — | — | — | — | — | — |
| | — | — | 27,989·70 | 26,000·31 | 24,037·42 | 22,101·03 | — |
| | — | — | — | — | (3,850) | (4,153) | (4,515) |
| 4 | — | — | — | — | — | — | — |
| | — | — | — | — | 26,040·47 | 24,104·08 | 22,194·19 |

lengths of the band heads are shown in parentheses in the first line of each row, the experimental wave-numbers of the zero lines in the second line, and the wave-numbers derived

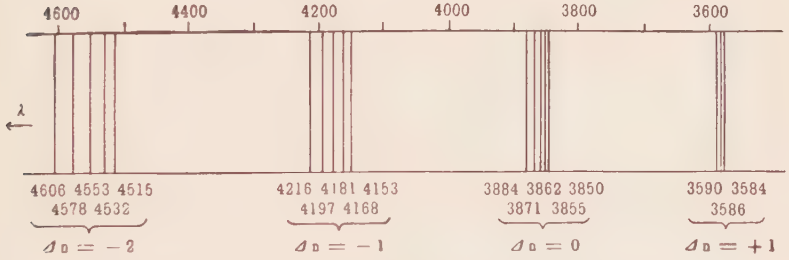
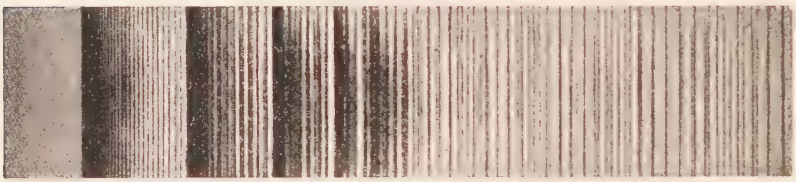



FIG. 115.—Arrangement of band heads in spectrum of $N_2$ (after Sommerfeld, "*Atomic Structure and Spectral Lines*").

from the formula in the third line. Fig. 115 shows the arrangement of the eighteen band heads contained in Table XXVIII, together with the relative wave-lengths. Four groups can be distinguished, corresponding to the quantum transitions $\Delta n$ equal to (− 2), (− 1), 0, and (+ 1). The group $\Delta n$ 0 is photographically reproduced in Fig. 116, in which the five band heads in the group are clearly visible.

It follows, of course, from the principle of term representation that, as has also been experimentally confirmed in many cases, the constants $a$ and $b$ (or $a'$ and $b'$) may be the *same* for *different* band systems of a molecule. The same amounts of molecular vibrational energy may be associated with different energy-levels in the electrons.

It was not until long after the bands themselves and the band heads had been investigated that the *electron terms* characteristic of molecules first became the object of experimental and theoretical research. In 1925 *Mulliken* [3] made the important discovery that, with respect to the electron terms, a far-reaching *analogy* exists *between molecular and atomic spectra*. Following Mulliken, we have to start from the hypothesis, which was first propounded by *Lewis*,[4] that those electrons in the atoms which are not combined



3884 3872 3862 3855 3850

FIG. 116. The so-called cyanogen bands in the nitrogen spectrum.

in closed groups, and are therefore in a sense supernumerary, join together in *common groups of eight*, the electrons then left over acting as the *optical electrons of the molecule*.

If, for example, we consider *diatomic molecules with thirteen electrons*, we should have to assume the presence of a closed $K$-group of two electrons in each of the two atoms, and, in addition, a common $L$-group of eight electrons, so that the thirteenth electron would remain over as an optical electron. The spectra of molecules with thirteen electrons must therefore be *analogous to the alkali spectra*. That such is actually the case has been demonstrated by *Mulliken*, *Birge*,[5] *Mecke*,[6] and other investigators with, for example, BeF, BO, CN: it has been found, moreover, that the spectra of diatomic molecules with twenty-one electrons, *e.g.*, of

[3] R. S. Mulliken, *Phys. Rev.*, **26** (1925), p. 561; *cf.* also *Phys. Rev.* (1925 *ff.*), *passim*.

[4] G. N. Lewis, *Trans. Faraday Soc.*, **19** (1923), p. 452.

[5] R. T. Birge, *Phys. Rev.* (1925 *ff.*) *passim*; *Nature*, **116** (1925), p. 783, and **117** (1926), p. 300.

[6] R. Mecke, *Zeitschr. f. Phys.*, **36** (1926), p. 795, and **42** (1927), p. 390.

MgF, AlO, SiN, exhibit relationships with the alkali spectra. Similarly, the molecular spectra of CO and $C^+O$ can be referred to the arc and the spark spectrum of magnesium, and so on.

The analogy between molecular and atomic spectra is especially marked in the *multiplicity of band spectra*. *Mecke*[7] was the first to show that molecular spectra also obey the so-called *alternation rule*, so that, with an even number of electrons in the molecule, the term multiplicity is odd, and conversely. On the basis of the multiplicity, Mulliken and Birge succeeded in establishing for the electron terms in molecules definitions entirely analogous to those holding for atomic spectra, so that, in band spectra also, we may speak for example of $^2S$- or $^3P$-terms. The rules for a systematic treatment of molecular electron terms have been developed principally by *Hund*.[8] It is, however, important to bear in mind that the multiplicity observed in band spectra is frequently an *effect of isotopy*.[9]

Like atoms, molecules as such may also be the agents in *fluorescence phenomena*. This effect has been particularly closely investigated by *Wood* [10] in the case of the *iodine molecule*. Wood irradiated iodine vapour at low pressure with mercury light of 5,461 Å. wave-length. It appeared that light-quanta of this wave-length are absorbed by the iodine molecule, so that the latter assumes an excited state of greater energy, from which it returns to the normal state with a *simultaneous* change in the nuclear vibrational and the electron energy. Corresponding to the circumstance that the rotational quantum number ($m$) is lowered or raised by one, numerous *doublets* appear in the fluorescence spectrum of iodine, Wood observing about twenty of them. Only on an increase of pressure or on an addition of an indifferent gas do the doublets develop into proper bands.

Among band spectra a particular significance attaches to those which are produced by diatomic *elementary molecules*, and, among these, the *molecular spectrum of hydrogen* occupies a special position. In the hydrogen molecule, as will be explained in the next section, the *moment of inertia*

[7] R. Mecke's first communication appeared in the *Zeitschr. f. Phys.*, **28** (1924), p. 261.

[8] F. Hund, *Zeitschr. f. Phys.*, **36** (1926), p. 657; *Physikal. Zeitschr.*, **28** (1927), p. 779.

[9] *Vide, e.g., R. Ritschl's* researches on the band spectra of the copper halides: *Zeitschr. f. Phys.*, **42** (1927), p. 172.

[10] R. W. Wood, *Phil. Mag.*, **35** (1918), p. 236. *Vide* also *W. Lenz's* interpretation in the *Physikal. Zeitschr.*, **21** (1920), p. 691.

is extraordinarily *small* as compared with that of other molecules. Consequently, as we can see from eqn. (1), the interval between adjacent lines in a band is proportionately large, so that the molecular spectrum of hydrogen does not resemble the usual band spectrum, but appears as a "*many-line spectrum*" stretching from the infra-red to the short-waved ultra-violet. It has, however, not yet been possible to represent the electron terms in this spectrum in the form of *Rydberg terms* except in a few instances;[11] greater success has been attained with the many-line spectrum of *helium*,[12] which may have its origin in transitory combinations between *metastable helium atoms*.[13]

A peculiarity of the band spectra of *elements* is an *alternating intensity* of the lines. With molecules consisting of *two similar atoms*, either *every second line* is missing from the bands, or the lines are alternately strong and weak. *Heisenberg*[14] correctly ascribed this alternation of intensity to a *resonance phenomenon of quantum mechanics*, referring it to his theory of the *arc spectrum of helium*. He had explained the characteristics of this atomic spectrum (*vide* § 107) by assuming that, in consequence of the resonance between the two planetary electrons, the system of energy-levels resolves into *two partial systems* between which *only weak intercombinations* are possible: these originate in the spin of the electrons and, in the absence of this spin, would not exist at all.

Now, according to Heisenberg, an exactly similar resonance phenomenon occurs in a molecule which contains *two atomic nuclei of the same structure*. The system of terms in such a molecule also resolves into two separate partial systems, between which intercombinations are only possible when a *resultant spin* is to be associated with each of the two nuclei. But, as Heisenberg was able to show, the lines in a band alternately originate in combinations *within* one and the same system of terms, and in *intercombinations*. Hence, when every second line is *completely* missing, it must be assumed that the atomic nuclei have *no* resultant spin, whereas a spin evidently exists when all the band lines can be observed, although alternating in intensity. The former condition holds for *helium* and *oxygen*, and the latter for

[11] *J. Sandemann* [*Proc. Roy. Soc.*, *A*. **110** (1926), p. 326] has, for example, identified terms of the form

$$18{,}567{\cdot}47 - R/(n + 0{\cdot}15988)^2 .$$

[12] *Cf.* W. E. Curtis and R. G. Long, *Proc. Roy. Soc.*, *A*. **108** (1925), p. 513.

[13] *Cf.* § 117.

[14] W. Heisenberg, *Zeitschr. f. Phys.*, **41** (1927), p. 239.

*hydrogen* and *nitrogen*. In point of fact, the number of nuclear electrons and protons is even for helium and oxygen, whereas there is an *odd* number of electrons in the nitrogen nucleus, and the hydrogen nucleus is identical with a proton.

### § 134. Nuclear Intervals and Dissociation Energy.

The band spectra of a diatomic molecule permit of the derivation of two quantities which are of importance in determining the structure of the molecule, *viz.*, the *nuclear interval* and the *dissociation energy*. In the first place, in accordance with § 133 eqn. (1), the molecular *moment of inertia* can readily be calculated from the structure of the bands, and, by § 132 eqn. (3), we have the relation

$$I = \mu r^2, \tag{1}$$

where $r$ is the distance between the two atomic nuclei and $\mu$ the so-called *reduced mass, i.e.*,

$$\mu = \frac{m_1 m_2}{m_1 + m_2}. \tag{2}$$

Hence, with a known chemical formula, the nuclear interval follows from the moment of inertia.

The *smallest moment of inertia* has been found to be that of the *hydrogen molecule* ($H_2$), which, in the normal state, has a value of $4{\cdot}67 \times 10^{-41}$ abs. units. The moment of inertia of the *iodine molecule* is especially large, two values of about 1,500 and 2,500 times that of hydrogen having been found. The moment of inertia must be particularly small for *hydrides* (compounds of hydrogen) in general, as follows from the formula for the reduced mass; for, with hydrides, this is always less than unity (when measured in atomic weights), whereas with nitrides and oxides it is not much smaller than 14 and 16 respectively [according to eqn. (2)].

The smallest value for the *nuclear distance* has likewise been found in the hydrogen molecule, *viz.* (in the normal state) 0·75 Å; a nuclear interval of 1·06 Å. corresponds to a second value of the moment of inertia for $H_2$. With the iodine molecule, the two values of the nuclear interval are 2·66 and 3·4 Å. By way of comparison, it will be remembered that, in Bohr's theory, the radius of the electronic orbit was found to be 0·53 Å.

It has been shown by *Mecke* [1] that, in chemically related compounds, *nuclear interval and chemical periodicity* are

[1] R. Mecke, *Zeitschr. f. Phys.*, **42** (1927), p. 390; *Naturwiss.*, **16** (1928), p. 521.

*closely connected.* This can be seen very clearly in Fig. 117, which represents the nuclear intervals (in Å.) of the *hydrides* as a *function of the atomic numbers* of the elements combined with hydrogen. The figure shows obvious *discontinuities* on passing from one period to another, while, within the separate periods, the curves exhibit a *uniform negative gradient.* The sudden rise at the commencement of a new period is evidently due to the addition of a new group of electrons, whereas the negative gradient of the curves may be referred to the increase in nuclear charge and consequent contraction of the atomic planetary systems. In the case of *nitrides*, *oxides*, and *elementary molecules*,[2] sharp discontinuities are likewise to be observed at the rare gases in

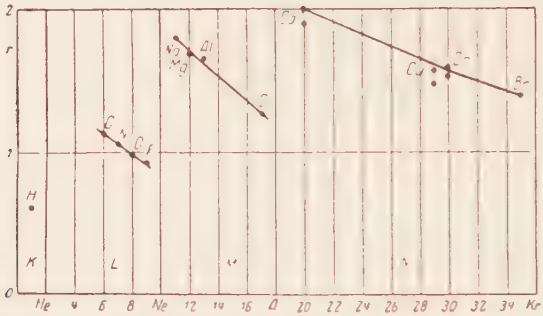



FIG. 117.—Nuclear intervals in hydrides (after Mecke).

the corresponding graphical representations, but, within the periods, the curves exhibit a minimum in the middle of the period, instead of a uniform gradient.

Furthermore, the *nuclear vibrational frequencies* in related compounds clearly evince a connection with chemical periodicity. The nuclear frequencies are especially *large* in *hydrides*, even in those of the higher elements. Comparatively small frequencies are observed with the diatomic molecules of iodine, potassium, and sodium.[3] In general the nuclear vibrational frequencies correspond to *infra-red wave-lengths* of an order of magnitude of between $10^{-4}$ and $10^{-2}$ cm. On the basis of the relations previously described, it is possible to *predict* approximately the nuclear vibrational frequency and the nuclear interval for any molecule, and conversely to derive these quantities from the experimental

[2] By elementary molecules we mean molecules composed of two similar atoms.

[3] As a rule, however, the vapours of the alkali metals are monatomic.

data of an observed band spectrum, when they are not already known.

In several band spectra *points of convergence* have been observed for the band heads, which, in long series, approach ever more closely together as the wave-length decreases, and finally *pass into a continuous absorption spectrum.* In those cases in which the convergence point cannot be directly observed, its position can be calculated by *extrapolation* from the wave-lengths of the observable members in the series. Since, in atomic spectra, a convergence point represents the ionization energy, the idea suggested itself to *Franck* [4] that, in a *band spectrum*, a *convergence point* denotes the *dissociation energy of the molecule.* The possibility thus arises of an *optical determination* of the *heat of dissociation*, which is generally derived from thermochemical observations.[5] In a calculation of this kind, however, account must be taken of the state of the atoms into which the molecule is resolved, whether it be normal, excited, or ionized. In the latter two cases the corresponding values of the energy have still to be deducted.

With the elementary molecules of *iodine*, *bromine*, and *chlorine*, for example, it has been found that one of the two atoms formed on resolution of the molecule is a normal one, whereas the other corresponds to the *metastable* $^2P_2$-state. The excess of energy inherent in this as compared with the normal ($^2P_1$-) state must therefore be subtracted from the calculated dissociation energy. A survey of the optical derivation of the dissociation heat of the above three

TABLE XXIX

*The Spectroscopic Determination of the Heats of Dissociation of Iodine, Bromine, and Chlorine*

| | Iodine | Bromine | Chlorine |
|---|---|---|---|
| Wave-length of the convergence point in Å | 4,995 | 5,107 | 4,785 |
| Equivalent potential in volts . . | 2·469 | 2·415 | 2·577 |
| Equivalent potential to be subtracted . | 0·937 | 0·454 | 0·109 |
| Dissociation potential in volts . . | 1·532 | 1·961 | 2·468 |
| ∴ Heat of dissociation (calcd.) in Kcal. | 35·2 | 45·2 | 56·5 |
| Heat of dissociation (obsd.) in Kcal. . | 34·5 | 46·2 | 57 |

[4] J. Franck, *Zeitschr. f. phys. Chem.*, **120** (1926), p. 144.

[5] *Cf.* the summary by H. Sponer in "*Ergebnisse der exakten Naturwissenschaften*" (Berlin; Springer, 1927), Vol. VI.

halogens, as carried out by *Kuhn*,[6] is contained in Table XXIX, which also shows the good agreement between the calculated values and those obtained from other sources.[7] As further examples of the optical determination of heats of dissociation we may mention $H_2$, with 100 Kcal. computed as against 97 observed ; $N_2$, with 274 and 263 respectively ; and $O_2$ with 162 and 150 respectively.

The dissociation of a molecule may be produced, not only by increasing the nuclear vibrations, but also by raising the *velocity* of its *rotation* until, on attaining a critical velocity, the molecule is *torn asunder* by the excessive *centrifugal force*. This possibility of dissociation finds spectroscopic expression in a sudden weakening and diffuseness of the lines in the bands as soon as a certain value of the rotational quantum number is reached, the lines shortly disappearing altogether. Thus *Ludloff* [8] has found that, with mercury hydride the series terminates at $m = 31$, with aluminium hydride at $m = 23$, and with calcium hydride at $m = 11$.[9]

Like the ionization of an atom, the dissociation of a molecule can be effected both by collisions with other molecules, and also by absorption of a light-quantum, *i.e.*, *photo-chemically*. In certain circumstances dissociation can also be produced by light of wave-lengths which are not absorbed by the reacting molecules, provided only that a *sensitizer* which can absorb the light be added. Owing to molecular collisions, the surplus energy absorbed is transferred from this sensitizer to the molecules of the substances concerned, which can then react. This appears especially clearly from an experiment of *Franck* and *Cario*, in which *mercury vapour* served as sensitizer for the *dissociation of hydrogen*.[10] For the resolution of a molecule of hydrogen into its two atoms a dissociation potential of 4·2 volts is required ; the corresponding wave-length is approximately 3,000 Å. A non-excited hydrogen atom, however, cannot absorb any radiation of greater wave-length than its resonance line, which (as the first line of the Lyman series) lies at 1,216 Å., far in the ultra-violet. Franck and Cario introduced mercury vapour into the hydrogen, and sub-

[6] H. Kuhn, *Zeitschr. f. Phys.*, **39** (1926), p. 77.

[7] The conversion of wave-lengths into volts follows § 93 eqn. (3), while 1 volt is equivalent to 23·07 kilogram-calories (Kcal.).

[8] H. Ludloff, *Zeitschr. f. Phys.*, **39** (1926), p. 528.

[9] The critical value of $m$ satisfies the condition that, for it, the second derivative with respect to $m$ of the expression representing the energy vanishes.

[10] G. Cario and J. Franck, *Zeitschr. f. Phys.*, **11** (1922), p. 161.

mitted the mixture at sufficiently low temperature to the action of the rays of 2,537 Å. wave-length emitted by a quartz mercury lamp. Although, in accordance with what we have just said, these rays could not be absorbed by the hydrogen, yet a dissociation of the hydrogen took place. This was clearly proved by a reduction of metallic oxides which, as is known from experience, can only be effected at the temperature in question by monatomic, but not by diatomic, hydrogen.

### § 135. Chemical Compounds.

Since the levels in the molecular vibrational energy can be read off from the band spectra, and since the corresponding nuclear interval can be calculated for each energy-level (with the aid of the moment of inertia), it has been found possible to represent the *potential energy of a molecular combination as a function of the nuclear interval.* The *minimum value* of the potential energy indicates the *dissociation energy*, which can thus be derived indirectly in this way also, when the course of the potential is known ; the corresponding abscissa represents the nuclear interval in the normal state. The occurrence of a *minimum* is due to the fact that the atoms themselves are composed of positive electrical nuclei and negative shells of electrons. Consequently, the attraction corresponding to negative energy values must become a *repulsion* (with *positive* energy values) for extremely small nuclear intervals, whereas, for an infinite nuclear distance, the energy of interaction vanishes.

Now the potential originates in forces of attraction on the one hand, and of repulsion on the other ; either the one or the other class of force may preponderate. Taking the forces as a whole, we may therefore resolve the *potential additively into two components*, one of which is always *negative* and due only to forces of *attraction*, whilst the other is always *positive* and arises only from forces of *repulsion*. As a first approximation, moreover, we may assume each component to decrease with a *power of the distance*. Hence we may put

$$V = -e^2 (C_1 r^{-p} - C_2 r^{-q}), \tag{1}$$

where the exponents $p$ and $q$ refer to the potential of the forces of attraction and repulsion respectively, and $C_1$ and $C_2$ are two constants. On the basis of these assumptions, the

actual course of the potential may, in fact, be approximately derived (Fig. 118). For hydrides $p$ is equal to one and $q$ to from three to four ; for oxides $p$ is found to be equal to from three to four, and $q$ to from six to nine. With the elementary molecules the exponents are still larger.[1]

Ever since the establishment of the electron theory it has seemed that all *chemical* processes should be referred to *electrical forces*. In 1916 *Kossel* [2] used this concept, in fact, as a basis for an electrostatic theory of valency, which, in some respects, had much in common with far older views that had been developed in 1812 by *Berzelius*. Kossel's theory led comparatively simply to the regarding of molecules as combinations between *oppositely charged ions*, *i.e.*, atoms which have either given up some of their own electrons, or have taken up additional electrons from outside, thereby losing their neutrality. Besides such "*heteropolar*" compounds of one positive and one negative ion, however, it was also necessary, at any rate, to assume combinations *between neutral atoms*, or so-called *homopolar compounds*. This assumption was unavoidable, if only to explain the existence of diatomic elementary molecules.

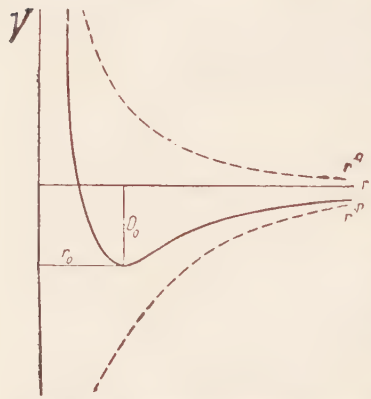


Fig. 118.—Course of potential in a molecular combination.

The experimental data furnished by *band spectra*, however, show quite clearly that the supposition of *electrostatic forces* is *insufficient* for the explanation of chemical compounds, and therefore the Coulomb forces must at any rate be supplemented by some other interaction. As was demonstrated by *Heitler* and *London* [3] in 1927, this effect arises in the *quantum-mechanical resonance* between the atoms forming a chemical compound.

Heitler and London began with an investigation into the interaction between *two neutral hydrogen atoms*. For this purpose they utilized *Schrödinger's equation* for the case of

[1] *Cf.* R. Mecke, *Zeitschr. f. Phys.*, **42** (1927), p. 390.
[2] W. Kossel, *Ann. der Phys.*, **49** (1916), p. 229.
[3] W. Heitler and F. London, *Zeitschr. f. Phys.*, **44** (1927), p. 455.

several particles [§ 97 eqn. (26)], expressing the potential energy in the form

$$(2) \quad V = e^2\left(\frac{1}{r_{ab}} + \frac{1}{r_{12}} - \frac{1}{r_{a1}} - \frac{1}{r_{a2}} - \frac{1}{r_{b1}} - \frac{1}{r_{b2}}\right),$$

where the subscripts $a$ and $b$ refer to the two nuclei, and 1 and 2 to the two electrons.

By applying the known solutions for the problem of the hydrogen atom, Heitler and London were able to calculate, by approximation, the *energy of interaction* between the two neutral hydrogen atoms. They obtained the result, at first sight surprising, that two such atoms can interact in *two quite different ways.* Schrödinger's equation, in fact, in its application to a system of two neutral hydrogen atoms, yields *two solutions* which, mathematically speaking, may be regarded as being relatively *symmetrical* and *anti-symmetrical.* In the second case, functions of the coordinates of the electrons change their sign on an interchange of the electrons, whereas, in the first case, an interchange is not accompanied by a change in sign.

Now, at moderate distances, the *symmetrical* solution leads to *negative* values for the energy of interaction, *i.e.*, to an *attraction* effect which weakens as the distance increases, but, at very close intervals, changes into a repulsion effect. The *anti-symmetrical* solution affords *only positive* values for that energy, *i.e.*, a *constant repulsion*, so that, when the two atoms approach each other very closely, *elastic reflection* must occur. The symmetrical solution, to which Heitler and London rightly looked for the explanation of *homopolar compounds*, also yields for the dissociation energy and nuclear interval values that satisfactorily agree with experiment. *Sugiura*,[4] who carried the approximations of Heitler and London's calculations still further, obtained 3·2 volts and 0·80 Å., and *Wang*,[5] using a similar method, 3·8 volts and 0·74 Å., whereas the values actually observed are 4·2 volts and 0·75 Å respectively.

Similar calculations to those applied to two neutral hydrogen atoms were also made by Heitler and London for two *neutral helium atoms* in the normal state. In this case *only* the *anti-symmetrical* solution was found to be possible, the symmetrical solution being excluded by *Pauli's principle.*

[4] Y. Sugiura, *Zeitschr. f. Phys.*, **45** (1927), p. 484.
[5] S. C. Wang, *Phys. Rev.*, **31** (1928), p. 579.

This principle therefore rules out a homopolar combination between two helium atoms in the normal state.[6]

Quite generally, Pauli's principle teaches that in an atom there may be *two*, but never more than two, electrons with *corresponding orbital motions*,[7] *i.e.*, not more than two electrons can have the same principal and subordinate quantum number, as well as the same orientation of orbital plane in a strong magnetic field.[8] If we regard two such corresponding electrons as an *electron pair*, the planetary systems of the atoms contain in general electron pairs, but, in some circumstances, *unpaired electrons* may also be present (the twin member, as it were, being lacking).

Now, as was recognized by *London*[9] in 1928, a *homopolar combination* may be regarded quite generally as a *pairing* of two corresponding electrons which *belong* as *unpaired* electrons *to different atoms*. The *chemical valency* of an atom is thus evidently determined by the *number* of its *unpaired electrons*, and this is obviously connected with the fact that the *valencies* of the elements in the second, fourth, sixth and eighth groups of the periodic system are *even*, whilst those of the other elements are *odd*.

The largest possible number of *mutually differing orbital motions* in an $n$-quantum group of electrons is, as was shown in § 103, $n^2$ (or half the number of electrons in the complete group); thus there are one in the one-quantum group, four in the two-quantum, nine in the three-quantum group, *etc.* On the other hand, as has been repeatedly explained, the "*outer*" group of electrons in all atoms, with the *highest quantum number*, contains one to eight electrons. The outer group is a two-quantum one in the second period of the system of the elements, a three-quantum group in the third period, and so on. Hence, *from the third period onwards, all the external electrons* may be *unpaired*. This is not the case for the *second period*, where, as we have just said, *only four* different orbits are possible for the outer electrons, so that, of *seven* outer electrons *only one* can be *unpaired*, and of *six only two*.

[6] A combination appears possible only between *excited* helium atoms; *cf.* § 117.

[7] We may also speak of translational motion, instead of orbital motion, in contrast to proper rotation, or spin.

[8] In the notation of § 103, we mean identity in the values of $n$, $l$, and $m$ ($l$). The concept of corresponding electrons is a narrower one than that of equivalent electrons in earlier sections. With equivalent electrons only $n$ and $l$, but with corresponding electrons all three quantum numbers governing the orbital motion must be identical.

[9] F. London, *Zeitschr. f. Phys.*, **46** (1928), p. 455, and **50** (1928), p. 24.

In agreement with chemical experience, London's theory shows that the *halogens* may have a valency of one, three, five, and seven, with the sole exception of *fluorine* which, belonging to the second period, can *only* be *univalent*—and actually no other valency has been detected for it. Likewise it follows from London's theory that the elements in the sixth group may have valencies of two, four, and six—with the exception of *oxygen*, which could *only* be *divalent* (and has, in fact, never been found with any other valency).

The valencies of an atom are wholly or partially *saturated* on entering into a molecular combination ; but it may, of course, happen that a molecule as a whole also may still contain unpaired electrons, in which case the unsaturated molecule appears as a so-called *radicle*. In no circumstances, however, can an electron that is already paired combine with a third ; an electron which has paired with a corresponding one in another atom takes no further part in the determination of valency.

London's theory of chemical compounds, moreover, leads directly to a *relation between valency and spectral multiplicity*. With $z$ reactive electrons, the quantum number of the resultant *electron spin* is equal to $z/2$, and hence (by § 101) the spectral multiplicity is $(z + 1)$. The valency is therefore equal to the *multiplicity diminished by one*. This rule proves to be universally obeyed, and it is noteworthy that the elements in the fifth group (the nitrogen group), in which ter- and quinquevalency are observed to the exclusion of univalency, exhibit quartet and sextet terms, but never doublet terms.

Finally, London [10] has shown how investigations on the reciprocal effects between two neutral atoms can be extended to the *interaction between three atoms*, and between atoms and molecules. It may also be mentioned in this connection that *Heisenberg* [11] has referred the phenomenon of *ferromagnetism* to a *quantum-mechanical resonance*.

The *chemical reactivity* of atoms and molecules depends essentially on whether the atoms or molecules are in an excited or in the normal state. The *absorption of light-quanta*, by raising the molecules to a higher-quantum state, may therefore originate, as *photochemical phenomena*, chemical processes which would not otherwise occur. In 1912 *Einstein* formulated the so-called *law of photochemical equivalence*, which states that the number of chemically

[10] F. London, "*Sommerfeld Festschrift*" (Leipzig ; Hirzel, 1928), p. 104.
[11] W. Heisenberg, *Zeitschr. f. Phys.*, **49** (1928), p. 619.

reacting molecules must be equal to the number of light-quanta absorbed, and this law is confirmed by experiment. Among well-known examples of photochemical reactions are the formation of hydrochloric acid (HCl) from $H_2$ and $Cl_2$, the assimilation of carbon dioxide by plants, and the decomposition of silver halides upon which photography is based. The dissociation of hydrogen in the presence of mercury vapour has already been mentioned as an example of a *sensitization* of a photochemical process. Another, very familar, example is the sensitizing of a photographic plate with suitable dyes which, such as eosin, possess in the red the absorption lines lacking, in that region, in the silver salts.

The converse of photochemical processes is represented by *chemico-luminescence*, due to a transformation of chemical energy into light-quanta. A well-known example of this is the oxidation of aldehydes at low temperatures, which is accompanied by an intensive luminescence. *Haber* and *Zisch* have shown that, on leading a stream of chlorine through sodium vapour, chemico-luminescence ensues; it can be recognized by the appearance of the characteristic *D*-lines of sodium at a temperature at which a luminescence due to heat is obviously excluded.[12]

Chemical reactions may be produced not only by the absorption of light-quanta, but also by *electron* and α-particle *collisions*. The experimental investigation of these processes promises to be of great importance, because it permits of an exact determination of the *minimum energy* requisite for the initiation of a chemical reaction.[13]

## § 136. The Atomic Lattice.

A special case of molecular formation is represented by the *atomic lattices* which we have already discussed in connection with the interference of X-rays. The *rock-salt crystal*, composed of sodium and chlorine ions, has been given as one example of an atomic lattice (*cf.* the earlier Fig. 97), and the structures of *diamond* and of *graphite* are reproduced in Figs. 119 and 120 as two further examples.[1] The problem before the *lattice theory of crystals*,[2] which has

[12] F. Haber and W. Zisch, *Zeitschr. f. Phys.*, **9** (1922), p. 302.

[13] *Cf.* especially S. C. Lind, "*The Chemical Effects of Alpha-Particles and Electrons*," 2nd edn. (New York; Chemical Catalog Co., 1927).

[1] In the diamond lattice, each atom represents the centre of a tetrahedron, the corners of which are occupied by four other carbon atoms. Graphite crystallises rhombohedrally.

[2] For further details of this subject *vide* especially M. Born's lectures, "*The Lattice Theory of Rigid Bodies*" ("*Problems of Atomic Dynamics*," Part II;

been developed principally by *Born*, *Kármán*, and *Debye*, is to deduce from the properties of the atoms, and from the forces acting between them in the lattice, the properties of the crystals constructed from the atoms, as well as their *macrophysical* behaviour, especially in *elastic*, *thermal*, *optical*, and *electrical* respects. The lattice theory thus establishes new *relations* between these branches of macrophysical phenomena, which are apparently fundamentally different, by referring them to the *same atomic-physical origins*.

Researches by *Madelung* and *Nernst* led to a first recognition of a close connection between the *coefficient o compressibility* (and other elastic constants) and the *infra-*

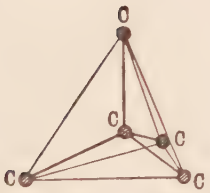



Fig. 119.—Structure of diamond.

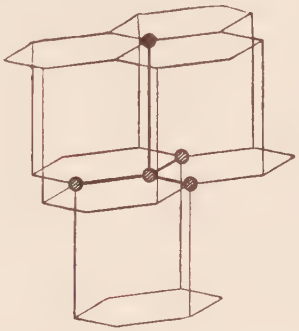


Fig. 120.—Structure of graphite.

*red residual rays* which characterize the various crystals in a spectroscopic respect. *Lindemann* and *Nernst* discovered a relationship between these infra-red wave-lengths and the *temperature of the melting-point*, on the basis of the assumption that, at the melting-point, the amplitude of the atomic vibrations becomes equal to the distance between neighbouring atoms. Nernst recognized also that the mechanical *cohesion* of a solid body must be essentially *identical* with the *chemical forces* at work in its interior.

On the grounds of the lattice theory, *Eucken* has explained the extraordinary *increase in the thermal conductivity* of crystals at low temperatures.[3] *Ewald* based on the lattice

Cambridge, Mass., 1926) and *K. F. Herzfeld*, "*Gittertheorie der festen Körper*" ("*Handbuch der Experimentalphysik*," Vol. VII; Leipzig, Akad. Verlagges., 1928).

[3] *Cf.* A. Eucken, *Physikal. Zeitschr.*, **29** (1928), p. 553. The conductivity of a quartz crystal at room temperatures is about five times that of quartz glass, whereas, at the temperature of liquid hydrogen, the corresponding ratio becomes about 500 to 1.

concept a theory of *double refraction* which is well supported in particular by the more recent researches on the refraction of X-rays. Furthermore, the behaviour of crystals in *electrical fields*, and the phenomena of *piezo-* and *pyro-electricity*, can all be derived from the assumption of electrical atomic forces.

*Hund* has made a successful attack, from the point of view of the lattice theory, on the fundamental problem of *chemical mineralogy*, *viz.*, the question why certain ionic species should be characterized by certain *types of lattice*. Hund starts with the assumption that the *most stable* structure will be that for which, with any given ions, the lattice energy is largest; and in this way he has succeeded in establishing a connection between the most important types already observed and the ionic species composing them. Finally, *Kossel* [4] has endeavoured to explain the *growth of crystals* with the help of this theory.

Conversely, of course, the lattice theory allows us to draw conclusions as to atomic properties from the experimental data of crystallography. The *determinations of ionic radii* by *Grimm*, *Herzfeld*, and *Fajans* [5] are of particular note in this respect. It was clearly demonstrated, for example, by comparing $O^{--}$, $F^{-}$, Ne, $Na^{+}$, and $Mg^{++}$, that an *electronic shell*, which appears as the outer shell in a rare gas, *contracts* as the nuclear charge increases, and *expands* as the charge decreases (*cf.* the earlier Fig. 117).

The *only solid bodies* regarded as *ideal* by modern atomic theory are *crystals*, *amorphous* solids being considered rather as exceedingly *viscous* liquids. It is, however, very remarkable that, from experiments on the diffraction of X-rays in certain liquids (especially alcohols), *Stewart* [6] has been led to the conclusion that even these *liquids* must be associated with a form of *lattice structure*.

## § 137. Electrolytes.

In some respects the theory of atomic lattices is closely related to a recent *theory of electrolytes*, which was developed in 1923 by *Debye* and *Hückel*.[1] Just as the individual ions in a polar crystal are surrounded by oppositely charged

[4] W. Kossel, *Physikal. Zeitschr.*, **29** (1928), p. 553.

[5] *Cf.* the article on the size and structure of molecules by K. F. Herzfeld and G. Grimm in the "*Handbuch der Physik*" (Berlin: Springer), Vol. XXII.

[6] G. W. Stewart and R. M. Morrow, *Phys. Rev.*, **29** (1927), p. 919.

[1] P. Debye and E. Hückel, *Physikal. Zeitschr.*, **24** (1923), pp. 185 and 305.

atoms, so (Debye assumes) every ion in a *strong electrolyte* is enveloped in an oppositely charged "*ionic cloud*" formed by the solvent. The existence of this cloud evidently produces a *decrease in conductivity*, because the force exerted on the cloud by an *external electric field* is *opposed* to that which is acting on the ion and driving it onwards. As *Debye* and *Falkenhagen* [2] recognized, however, this so-called electrophoretic impeding force is supplemented by a second obstructive force, which they called the *relaxation force*. This originates in the necessity of a certain finite time for the development of the ionic cloud, and this time is characterized by a constant (the so-called relaxation time). Since the cloud does not follow the movement of the ion, so that the latter has to form new clouds all along its path, it is easy to understand that the cloud of a *moving* ion will, as regards its *distribution of charge*, exhibit a *dissymmetry* which impedes the motion. If we suppose the cloud to be divided into two halves by a plane normal to the direction of motion, the charge on the forward half will always be the smaller.

Now, in *very strong fields*, the velocity of the ions will be so large that, as it were, there will be no time for the development of ionic clouds. Since the impeding forces then become insignificant, we may expect an *increase in the conductivity with the potential*, when a sufficiently large intensity of field is attained. This phenomenon has actually been observed by *Max Wien*.[3] Thus the so-called *Wien effect* is referred by Debye and Falkenhagen's theory of electrolytes to the relaxation of the ionic clouds.

### § 138. The Polarity of Molecules.

By introducing the concept of ionic clouds into the theory of electrolytes, *Debye* made a considerable advance, as he also did when he evolved the hypothesis of *polar molecules* (so-called *dipole-molecules*) in the theory of *non-conductors*.[1] He began by assuming that molecules with an *asymmetrical distribution of charge* must behave like *elementary magnets*. They must therefore take up such an *orientation* in a homogeneous *electric field* that the axis of the moment lies in the direction of the field, but, as in the case of paramagnetism (*cf.* § 116), this attempt at orientation

[2] P. Debye and H. Falkenhagen, *Physikal. Zeitschr.*, **29** (1928), p. 121.

[3] M. Wien, *Ann. der Phys.*, **83** (1927), p. 327, and **85** (1928), p. 795.

[1] *Vide* P. Debye, "*Polar Molecules*" (New York; Chemical Catalog Co., *in the press*).

will be opposed by *thermal motion*, so that the *moment* produced in a volume unit must *depend on the temperature*.

On the other hand, however, as physicists have long known, an electric field creates an electric moment in the molecules of a non-conductor also by "*deformation*," *viz.*, by small *displacements* of the electrically charged particles. Quite generally, therefore, the *moment generated* is made up of *two components*: one component is a deformation effect and independent of the temperature, whilst the other is an orientation effect and varies with the temperature. The reaction of the created moment on the field produces, as is well known (*e.g.*, with the plate condenser), those phenomena which *Faraday* had interpreted in terms of the *dielectric* and the dielectric constant.

Thus, according to Debye's theory, the *dielectric constant* must *depend on the temperature*, in the case of *polar molecules*. Conversely, when this dependence is experimentally observed we may assume a polar character in the molecules concerned, and we can calculate the *value of the polar moment* from the *known temperature variation*. The values so derived are of the order $10^{-18}$ cm. $\times$ E.S.U., corresponding to a value of $4{\cdot}77 \times 10^{-10}$ E.S.U. for the elementary charge, and of about $10^{-8}$ cm. for the molecular intervals.

Experiment also confirms the expectation that a *polar character* will *not* be associated with molecules with a *symmetrical structure*, chemically speaking, but only with *asymmetric* molecules. Great significance attaches in this respect to observations made by *Sänger*,[2] who investigated the series of compounds $CH_4$, $CH_3Cl$, $CH_2Cl_2$, $CHCl_3$, and $CCl_4$ in vapour form at constant density. He found that, in those compounds which contain *only hydrogen* or *only chlorine atoms* besides the carbon atom, the dielectric constant is actually *independent* of the temperature, whereas, with the three other molecules, he proved that it varied with the temperature. Finally, an analogy to Stern and Gerlach's researches previously mentioned (§ 116) was furnished by *Wrede*,[3] who demonstrated an experimental *resolution of polar molecular rays in an electric field*.

## § 139. Metals.

The solid state of aggregation includes, as a special case, the *metallic state*. The idea has suggested itself of explaining the characteristics of this state, especially the high thermal

[2] R. Sänger, *Physikal. Zeitschr.*, **27** (1926), p. 556.
[3] E. Wrede, *Zeitschr. f. Phys.*, **44** (1927), p. 261.

and electrical conductivities, by assuming within metals the existence of *free electrons* which behave as a *gas*.[1] This concept, however, was faced with difficulties that were formerly regarded as insuperable, for it followed from the kinetic theory of gases that, at a given temperature, the kinetic energy of every free electron must be equal to that of a monatomic gas molecule. It would therefore be expected that *free electrons* would *noticeably contribute towards the specific heat of metals*—a conclusion that is obviously *contradicted by experience*.

As was shown in 1927 by *Sommerfeld*,[2] however, this difficulty disappears if, in dealing with the electrons, we replace classical by a new form of *statistics*, which was developed directly from *Pauli's principle* by *Fermi*,[3] in 1926. Fermi's statistics [4] is based on the hypothesis that, corresponding to the three degrees of freedom of its *translational motion*, every *molecule of a gas* is to be associated with *three quantum numbers*, but that, in a gas, there can *never* be *two molecules* for which *all the quantum numbers* are *identical*. If $\nu$ be any fundamental frequency [5] characteristic of the gas, and if we suppose an energy element $\epsilon$ of the gas to be defined by the product $h\,\nu$, it may be deduced from Fermi's statistics that there cannot be more than $(n + 1)\,(n + 2)/2$ molecules of energy $n\,\epsilon$. Thus, there can be only one molecule with zero energy, only three molecules with an energy $\epsilon$, only six with an energy $2\,\epsilon$, only ten with an energy $3\,\epsilon$, *etc.*, whilst energy values other than integral multiples of $\epsilon$ cannot occur at all.

Contrary to the classical system, Fermi's statistics leads to the conclusion that, at the *absolute zero of temperature*, *neither the energy nor the pressure of a gas vanish entirely*, and, furthermore (again in opposition to classical theory), that the *specific heat* at sufficiently low temperatures is *proportional* both to the *temperature* and to the *mass of the gas particles*. It follows also from Fermi's statistics that the "degeneration" of a gas is determined by the product of the molecular mass and the absolute temperature, and

[1] The electron theory of metals was initiated by *Drude* in 1900, and was developed principally by *H. A. Lorentz*.

[2] A. Sommerfeld, *Zeitschr. f. Phys.*, **47** (1928), p. 1; *Naturwiss.*, **15** (1927), p. 825.

[3] E. Fermi, *Zeitschr. f. Phys.*, **36** (1926), p. 902.

[4] *Cf.* the author's "*Wave Mechanics and the New Quantum Theory*" (London, Constable; New York, van Nostrand, 1928), chap. xiv.

[5] In the case of a gas contained in a vessel, a fundamental frequency of this kind is determined by the size of the vessel; if this be cubical in shape, by the length of one side.

hence that, the smaller the mass of the gas particles, the higher is the temperature at which degeneration of the gas commences. Since the *mass of an electron* is about 1,800 times as *small* as the mass of the lightest atom, it is a necessary consequence of Fermi's theory that, even at *room temperatures*, the *electronic gas* in a metal is *completely degenerate*. Hence, on account of the proportionality then existing between specific heat and mass, the contribution of the free electrons towards the specific heat of the metal would actually be devoid of all significance.[6]

It has been demonstrated by Sommerfeld that the application of Fermi's statistics to the electronic gas leads to the well-known *law* of *Wiedemann* and *Franz*, according to which the *ratio* between the *electrical* and the *thermal conductivity* of a metal is universally *proportional to the absolute temperature*. The value yielded by Sommerfeld's theory for the proportionality factor is given by

$$a = \frac{\pi^2}{3}\left(\frac{k}{e}\right)^2, \tag{1}$$

where $k$ denotes Boltzmann's constant (the gas constant divided by the Loschmidt number), and $e$ the elementary quantum of electricity. The older theory of *Drude* led to the same expression, but with a numerical factor of 3 instead of $\pi^2/3$, whilst the later, more accurate, theory of *H. A. Lorentz* gave a factor of 2. At a temperature of 18° C., the ratio between electrical and thermal conductivity is 6·3 according to Drude, 4·2 according to Lorentz, and 7·1 according to Sommerfeld. The experimental mean from observations on twelve different metals is 7·11.

Sommerfeld has also successfully applied Fermi's statistics to the theories of *contact potentials*, of the *Thomson effect*, and of the *Peltier effect*.[7] As he showed, the magnitude and the variation with temperature of these effects are correctly accounted for by Fermi's statistics. It is to be regarded as a great merit of Sommerfeld's theory of metals that, in contradistinction to the older theories, special hypotheses

[6] Whereas, according to classical statistics, the contribution of the free electrons to the molecular heat of the metal would be equal to 1·5 times the gas constant, Fermi's statistics shows that, at room temperatures, this amount would be equal to only one-hundredth of the gas constant.

[7] The Peltier effect consists in the heating of one joint and the cooling of the other joint observed in an electrical circuit composed of two different metals. The Thomson effect appears essentially as a transport of heat by an electric current.

on the interaction between the free electrons and the atomic residues of the metal are rendered superfluous.[8]

[8] We have, perhaps, to regard as a still unexplained quantum effect certain phenomena discovered by *Kamerlingh Onnes*. On reaching certain extremely low temperatures, *some* metals (chiefly lead, mercury, tin, and thallium) pass *suddenly* into a state of so-called *supraconductivity*, in which no electrical resistance at all can be detected. The transition temperature lies, for example, for lead at 7·2°, for mercury at 4·15°, for tin at 3·78°, and for thallium at 2·32° above the absolute zero. *Vide* the summary by C. A. Crommelin, *Physikal. Zeitschr.*, **21** (1920), pp. 274, 300, 331.

## Chapter XVII

# THE INTERACTION BETWEEN LIGHT AND MATTER

### § 140. Dispersion.

The development of wave and quantum mechanics formed the first possible basis for an exact treatment of those phenomena which originate in an *interaction between electromagnetic waves and atoms*. Of such phenomena the longest known is the *dispersion of light*, discovered by Newton. *Schrödinger* [1] has established a *wave-mechanical theory of dispersion*, by applying his fundamental equation (in a more general form, which takes into account also a possible dependence of the field-scalar on the time [2]) to the *behaviour of an atom in an alternating electric field :* a light-wave must, indeed, be regarded as an oscillating electric field. If the square of the magnitude of the *field-scalar* be interpreted as the *density of electrical charge*, the equation for the field-scalar yields an expression for the *accessory electric moment* which is generated in the atom by the alternating field, and is itself rapidly oscillating in character. The vibrations of the accessory moment, however, produce a *secondary radiation* which is superimposed on the primary radiation ; and hence the velocity of propagation of the light-wave varies in a degree depending on the frequency, so that dispersion occurs.

Let us suppose the states of the atom permitted by the quantum theory to be numbered, and the atom to be in the $k$th state at the instant when the light-wave strikes it. Then, if $E_k$ be the energy associated with this state, $F$ the amplitude of the electric field-strength, and $\nu$ the primary frequency, Schrödinger's theory of dispersion leads to the following expression for the accessory electric moment :—

$$(1) \qquad M = a_{kk} + 2\,F\cos 2\pi\nu t\,.\sum_n \frac{(E_n - E_k)\,a_{kn}\,b_{kn}}{(E_n - E_k)^2 - h^2\nu^2}\,.$$

[1] E. Schrödinger, *Ann. der Phys.*, **81** (1926), p. 139 ; "*Collected Papers on Wave Mechanics*" (London ; Blackie, 1928).

[2] The extended form of Schrödinger's equation, a derivation of which must be omitted here, reads :

$$\nabla^2 S - \frac{4\pi i m}{h}\frac{\partial S}{\partial t} - \frac{8\pi^2 m}{h^2} V S = 0.$$

In this equation, the quantity $a_{kk}$ determines the probability of the duration in the $k$th state ; it is thus independent of the time, and hence does not affect the dispersion. The quantities $a_{kn}$ and $b_{kn}$ are related to the characteristic functions of the atom in a way which we cannot discuss here ; they determine essentially the *probability of a spontaneous transition* between the $k$th and $n$th states.

As eqn. (1) shows us, the phenomena of dispersion depend in no way on the actual vibrational or orbital frequences of the electrons in the atoms, but on the *possible spectral frequencies* of the atom on which the light-wave impinges. Moreover, they depend not on any individual, but on *all* possible frequencies *collectively*, although these may not actually occur. *Ladenburg* [3] (1921) was the first to recognize this important fact. His dispersion formula, which was indeed still incomplete, was further developed by *Kramers*,[4] who introduced the so-called *negative terms* ($n < k$), corresponding to emissive processes, whereas Ladenburg had taken into account only terms for which $n > k$, corresponding to absorptive processes.[5]

The new dispersion formula, as was shown by *Kramers* and *Heisenberg*,[6] leads to the important conclusion that secondary radiation may also contain light of a *combinational frequency*, *i.e.*, of a frequency equal to the *sum* or the *difference* of the primary frequency and a *characteristic* frequency of the dispersing atom or molecule. This theoretical deduction has apparently been empirically confirmed by the discovery in 1928 of the *Raman effect.*

Raman [7] found that, when monochromatic light is passed through various liquids, the *spectrum of the scattered radiation* contains sharp lines, the frequencies of which are both larger and smaller than the primary frequency. The *frequency difference*, however, proved to be *independent of the primary frequency*, and, in the case of *benzene*, Raman was also able to demonstrate that one of the frequency differences coincided with an *infra-red characteristic frequency of the*

[3] R. Ladenburg, *Zeitschr. f. Phys.*, **4** (1921), p. 451.

[4] H. A. Kramers, *Nature*, **113** (1924), p. 673.

[5] When the primary frequency coincides with a spectral frequency of the atom, *resonance* occurs. In apparent contradiction to the dispersion formula [eqn. (1) ], however, the resonance energy does not become infinite, because correction terms disregarded in deriving the formula, which is only approximately valid, cannot be neglected in the case of complete resonance.

[6] H. A. Kramers and W. Heisenberg, *Zeitschr. f. Phys.*, **31** (1925), p. 681. *Cf.* an earlier discussion by A. Smekal, *Naturwiss.*, **11** (1923), p. 873.

[7] C. V. Raman and H. P. Krishnan, *Nature*, **121** (1928), p. 501 ; *Indian Journ. Phys.*, **2** (1928), pp. 1 and 309.

*benzene molecule*, known from other sources. We may therefore, perhaps, justifiably hope that, with the aid of a single Raman photograph, it may be possible to derive the whole of the infra-red frequencies characteristic of a molecule, which are much more difficult to investigate in any other way.

### § 141. The Photo-Electric Effect.

It follows from the theory of differential equations that, in general, the whole *region of characteristic values* in a problem of atomic mechanics resolves into a *discontinuous* and a *continuous partial region*. In the especially simple case of the *hydrogen atom*, the discontinuous region connotes *negative energy* and a *closed, periodic movement* (*i.e.*, elliptical orbits), whereas the continuous region means *positive energy* and an *aperiodic motion* (*i.e.*, hyperbolic orbits). The boundary between the two regions is represented by the term corresponding to the *ionization energy* of the hydrogen atom. Analogous, but more complicated, relations hold for other atoms.[1]

If, now, an electron in an atom is raised by external action to a higher level which belongs to the region of continuous characteristic values, its previously periodic motion becomes *aperiodic*, and the electron *deserts* the atom. When the liberation from the atom is brought about by a light-quantum, we have to deal with the well-known *photo-electric* or *photo-effect*.

Thus, as was recognized by *Wentzel* [2] and by *Beck*,[3] the photo-electric effect may be referred to the wave-mechanical theory of dispersion. When the "disturbed" characteristic value falls within the continuous region, *Schrödinger's dispersion equation* does, in fact, describe the propagation of a *material wave* emanating from the atom on which the light falls. This wave determines, in the sense of *Born's statistical interpretation* (previously discussed [4]), both the *probability of the emission of a photo-electron* and the dependence of this probability on the direction.

Theoretical deductions lead to the experimentally confirmed result that the emission of electrons in a given direction is proportional to $\cos^2\theta$, where $\theta$ is the angle

[1] In general, the discontinuous and continuous zones may, of course, overlap.

[2] G. Wentzel, *Zeitschr. f. Phys.*, **40** (1926), p. 574; **41** (1927), p. 828.

[3] G. Beck, *Zeitschr. f. Phys.*, **41** (1927), p. 443.

[4] *Vide* § 99.

included between this direction and that of the *electric vector in the light-wave ;* the *maximum* intensity therefore coincides with the direction of this vector. The *yield* of the photo-effect, or the number of electrons liberated in unit-time, as derived from the wave-mechanical dispersion theory, is also confirmed by experiment. Furthermore, theory indicates that the yield depends, not only on the frequency and the atomic number, but also on the sub-ordinate quantum number of the sub-group from which the electron is expelled.

A special variety of the photo-electric effect is represented by the so-called *Auger-effect*,[5] discovered in 1925. This phenomena arises when, by the action of X-rays, an inner electron is torn away from an atom of the irradiated substance, and the position thus vacated is then filled by an electron from a group of higher quantum number. The energy liberated in this process may either be transformed into a light-quantum, in which case X-rays are emitted as a *secondary* or *fluorescent radiation*, or else, in some circumstances, it may serve to remove a further, so-called Auger, electron from the atom. The latter electron leaves the atom with a kinetic energy equal to the difference between the liberated energy and its own binding energy.

*Wentzel* [6] and *Fues* [7] have treated the Auger-effect from the point of view of *wave-mechanics* by investigating the material wave emitted by an atom from which an inner electron has been expelled. The statistical interpretation of this material wave apparently determines the probability of the expulsion of an Auger electron. On the other hand, however, we know the probability of the emission of a light-quantum consequent on the quantum transition occurring within the atom.[8] Hence we can calculate the proportion of *quantum transitions accompanied* to those *unaccompanied by radiation*, or, in other words, the *yield in fluorescence* of the absorbed X-rays. The theoretical results agree well with Auger's experiments.

[5] P. Auger, *Journ. de Phys.*, **6** (1925), p. 25 ; **6** (1926), p. 183.
[6] G. Wentzel, *Zeitschr. f. Phys.*, **43** (1927), p. 524.
[7] E. Fues, *Zeitschr. f. Phys.*, **43** (1927), p. 726.
[8] This quantum transition refers, of course, to the transition of the electron from the higher quantum orbit into the vacated position of lower quantum number.

## § 142. The Transformation of Electronic Radiation into Light-Quanta.

Since, in the photo-electric effect, atoms emit electrons when acted upon by light waves, the *emission of light-quanta by atoms under the action of material waves* may be regarded as a converse of the photo-effect. A process of this kind underlies the *transformation of arrested cathode-rays into X-rays*, which takes place in *X-ray tubes*.

In this phenomenon also, the *light-quantum relation* is found to be most accurately obeyed. The largest producible light-frequency is that which, when multiplied by the elementary quantum of action, is equal to the product of the electron charge and the tube potential. The *continuous X-ray radiation* generated in a tube must therefore exhibit, on the *short wave-length side*, a *sharp boundary* which, as the tube potential is increased, will be *displaced* in the direction of smaller wave-lengths, the displacement being governed by the universal relation between wave-length and potential [§ 93 eqn. (3)]. How accurately this relation is satisfied has been demonstrated by measurements of *Duane* and *Hunt*,[1] extending over a range from a few thousand to about 200,000 volts. The accuracy is, indeed, so great, that this displacement rule has enabled *Duane* and *Blake* [2] to effect an exact determination of the *elementary quantum of action*.

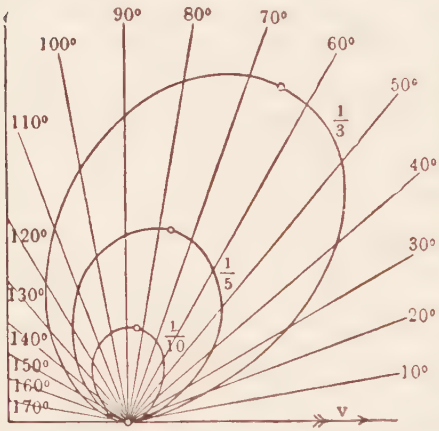



FIG. 121.—Intensity distribution in X-rays generated by cathode rays (after Sommerfeld).

The formulæ describing the *angular variation of the radiation intensity*, which were first accurately derived by *Sommerfeld*,[3] are very complicated. Fig. 121 is a reproduction of the intensity distributions, theoretically calculated by Sommerfeld, for cathode rays with velocities of $c/10$,

[1] W. Duane and F. Hunt, *Phys. Rev.*, **6** (1915), p. 166.
[2] W. Duane and F. C. Blake, *Phys. Rev.*, **10** (1917), pp. 93 and 624.
[3] *Cf.* A. Sommerfeld, "*Atomic Structure and Spectral Lines.*"

$c/5$, and $c/3$. The curves show a *dissymmetry* increasing with the velocity.

The processes in X-ray tubes find essentially a parallel in the *flashing out of optical spectral lines*, to be observed in experiments on *electron impacts*. The phenomenon of electron impacts has been treated from the point of view of *wave-mechanics* by *Born*.[4] The solution obtained by him not only describes correctly elastic and inelastic reflections (*vide* § 93), as well as ionization by electron collisions, but it also allows of a calculation of the *yield* of elastic and inelastic impacts.

### § 143. The Compton Effect.

When a *light-quantum collides* with a *material particle*, it follows from the corpuscular nature of light that the collision leads to not only a *change in direction*, but also an *alteration in frequency* of the light-quantum. Let us denote by $m_0$ the stationary mass of the material particle, by $\mathbf{v}$ the vector of its velocity, by $\nu$ the frequency, and by $\mathbf{f}$ a unit-vector with the direction of the ray. We shall further put, for brevity,

$$\beta = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, \tag{1}$$

and distinguish with dashes the quantities *after* the collision. According to the Laws of the *Conservation of Energy* and of the *Conservation of Momentum*, the two following relations must then be satisfied :—

$$h\nu + \beta m_0 c^2 = h\nu' + \beta' m_0 c^2, \tag{2}$$

and

$$\frac{h\nu}{c}\mathbf{f} + \beta m_0 \mathbf{v} = \frac{h\nu'}{c}\mathbf{f}' + \beta' m_0 \mathbf{v}'. \tag{3}$$

These two equations are, of course, *not* sufficient for the determination of the quantities $\nu'$, $\mathbf{f}'$, and $\mathbf{v}'$ from given values of $\nu$, $\mathbf{f}$, and $\mathbf{v}$ ; but they do enable us to calculate the *change in frequency* for a *given angle of scatter*, *i.e.*, for a given alteration in direction of the light-quantum.

In the following computation we shall consider the especially simple case in which the material particle is *at rest* before the collision, so that $v = 0$ and $\beta = 1$ ; the impact may then at any rate be regarded as *plane*. Let $\varphi$ be the

[4] M. Born, *Zeitschr. f. Phys.*, **37** (1926), p. 863 ; **38** (1926), p. 803.

angle of scatter, or the angle included between the directions of the primary and secondary rays, and let $\psi$ be the angle included between the direction of the primary ray and that

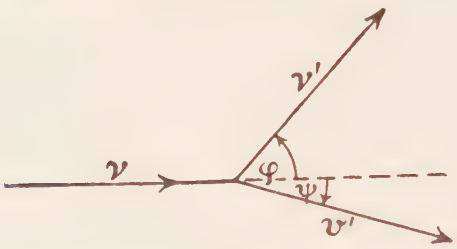



Fig. 122.

of the motion of the particle after the collision, $\psi$ being taken in the opposite sense to $\varphi$ (Fig. 122). The energy equation (2) then assumes the form

$$h\nu + m_0 c^2 = h\nu' + \beta' m_0 c^2 ; \tag{4}$$

whilst the vectorial momentum equation may be resolved into the two equations

$$\frac{h\nu}{c} = \frac{h\nu'}{c}\cos\varphi + \beta' m_0 v' \cos\psi , \tag{5}$$

and

$$0 = \frac{h\nu'}{c}\sin\varphi - \beta' m_0 v' \sin\psi . \tag{6}$$

In order to derive, for a given angle $\varphi$, the altered frequency $\nu'$ from these three equations, it is simplest to set to work as follows. Regarding the material particle as being an *electron*, so that $m_0$ denotes the mass of an electron, we introduce a *universal frequency* $\nu_0$ as defined by the relation

$$h\nu_0 = m_0 c^2 . \tag{7}$$

On dividing then eqn. (4) by $h\nu$ and squaring, we obtain

$$\left(1 + \frac{\nu_0}{\nu} - \frac{\nu'}{\nu}\right)^2 = \beta'^2 \left(\frac{\nu_0}{\nu}\right)^2 . \tag{8}$$

Again, we find from eqn. (5)

$$\beta' v' \nu_0 \cos\psi = c(\nu - \nu'\cos\varphi) , \tag{9}$$

and from eqn. (6)

$$\beta' v' \nu_0 \sin\psi = c\nu' \sin\varphi . \tag{10}$$

On squaring these two equations, adding together, and finally dividing by $c^2 \nu^2$, we obtain

$$\frac{\beta'^2 v'^2 \nu_0^2}{c^2 \nu^2} = 1 + \left(\frac{\nu'}{\nu}\right)^2 - 2\frac{\nu'}{\nu}\cos\varphi . \tag{11}$$

But, by eqn. (1)

$$\frac{v'^2}{c^2} = 1 - \left(1 - \frac{v'^2}{c^2}\right) = 1 - \frac{1}{\beta'^2} . \tag{12}$$

Hence eqn. (11) may be written in the form

$$1 + \left(\frac{\nu'}{\nu}\right)^2 - 2\frac{\nu'}{\nu}\cos\varphi + \left(\frac{\nu_0}{\nu}\right)^2 = \beta'^2\left(\frac{\nu_0}{\nu}\right)^2 . \tag{13}$$

The left-hand sides of eqns. (8) and (13) must therefore be equal, and a simple calculation gives us

$$\frac{\nu_0}{\nu} - \frac{\nu'}{\nu} - \frac{\nu_0}{\nu}\frac{\nu'}{\nu} = -\frac{\nu'}{\nu}\cos\varphi ,$$

or [1]

$$\frac{\nu'}{\nu} = \frac{1}{1 + 2\dfrac{\nu}{\nu_0}\sin^2\dfrac{\varphi}{2}} . \tag{14}$$

Thus

$$\frac{\nu}{\nu'} - 1 = 2\frac{\nu}{\nu_0}\sin^2\frac{\varphi}{2} .$$

If, now, we put

$$\nu' = \nu + \Delta\nu ,$$

we have

$$\frac{\nu}{\nu'} = 1 - \frac{\Delta\nu}{\nu'} ,$$

and hence

$$\Delta\nu = -2\frac{\nu\nu'}{\nu_0}\sin^2\frac{\varphi}{2} . \tag{15}$$

When $\Delta\nu$ is small as compared with $\nu$, we may replace $\nu\nu'$ by $\nu^2$. Since, in accordance with § 111 eqn. (6),

$$\Delta\lambda = -\frac{c}{\nu^2}\Delta\nu ,$$

and we may put

$$\frac{c}{\nu_0} = \lambda_0 , \tag{16}$$

[1] Using the well-known identity

$$1 - \cos\varphi = 2\sin^2\frac{\varphi}{2} .$$

we obtain finally

$$\Delta\lambda = 2\,\lambda_0 \sin^2 \frac{\varphi}{2}\,. \tag{17}$$

The value of the universal wave-length $\lambda_0$ follows at once from eqn. (7) :

$$\lambda_0 = 0{\cdot}0243 \text{ Å}. \tag{18}$$

The alteration in wave-length therefore depends only on the angle of scatter, and proves to be *independent of the wave-length* itself. Hence, the *smaller* the *wave-length* of the scattered light, the larger is the *relative change in wave-length*. With visible light, the alteration as determined by eqn. (17) would amount to only about $\frac{1}{100\,000}$th of the wave-length. On the other hand, a noticeable effect may be expected with *X-rays*, the wave-lengths of which are of the order 1 Å. to 0·1 Å.

In point of fact, *Arthur H. Compton* [2] was able in 1922 to show that X-rays experience the *decrease in frequency*, demanded by theory, on scattering by light atoms. Compton's first successful experiments were made with the $K_\alpha$-radiation of molybdenum (0·71 Å). The scattering agent was graphite, *i.e.*, carbon atoms ; the increase in wave-length for normal scatter was found to be 0·044 Å., in almost complete agreement with theory. The existence of the relation quoted in eqn. (17) between line displacement and direction of scatter was confirmed in researches by *Compton* and *Ross*. They showed also that, in the scatter spectrum, the displaced line is always accompanied by the *undisplaced*, the strength of which is the greater, the higher the atomic weight of the scattering substance.

The so-called *recoil electrons*, which theoretically must form a companion phenomenon to scattering, have been investigated principally by *Compton* and *Hubbard*.[3] Using two needle counters, *Bothe* and *Geiger* [4] registered the recoil electrons on the one hand, and on the other, by means of their ionizing effect, the scattered light-quanta ; and in this way were able clearly to detect *coincidences* between the respective throws of the needles. Thus they established the important fact that the phenomenon of electron recoil is *linked*, not only statistically, but also *individually* with the

[2] *Vide* the comprehensive treatment in A. H. Compton's "*X-rays and Electrons*" (New York : van Nostrand, 1927).

[3] A. H. Compton and J. C. Hubbard, *Phys. Rev.*, 23 (1924), p. 439.

[4] W. Bothe and H. Geiger, *Zeitschr. f. Phys.*, 32 (1925), p. 639.

scattering process, for a *single recoil electron* is produced on the *scattering of each single light-quantum*. Experiments by *Compton* and *Simon* [5] have actually proved that, in each individual case, the relation demanded by eqns. (5) and (6) exists between the directions of recoil and of scatter. Although it has frequently been questioned in recent years whether the laws of nature are, perhaps, only statistical laws, observations on the Compton effect show that no mere statistical interpretation is to be attached to the Principles of the Conservation of Energy and of Momentum, but that the impact between a light-quantum and an electron is governed by them in *each individual case*.

It is only the angle of scatter that is arbitrary and a matter of chance in each separate instance, and hence the question as to the *dependence of the intensity of the scattered radiation upon the angle of scatter* does actually appear as a statistical problem, which could only be accurately solved on applying the concept of *material waves*. The Compton effect was first treated from the point of view of the new quantum mechanics by *Dirac*,[6] and subsequently, along the lines of Schrödinger's method, by *Gordon*.[7] Both these investigators considered the problem of a *free* electron under the influence of electromagnetic waves, and they obtained for the intensity distribution relations which agreed better with experiment than those previously derived.

A more perfect theory of the Compton effect was developed by *Wentzel*,[8] who regarded the scattering electrons no longer as free, but as *associated with the atom*. In this way Wentzel came to the conclusion that the process of scattering must yield a *continuous spectrum ;* but, with the help of Schrödinger's dispersion theory, he was able to show that this spectrum must exhibit a sharply defined *intensity maximum*, when the effect of the light-quantum on the scattering electron preponderates over the attraction of the corresponding atomic nucleus. The intensity maximum represents the observed displaced line, and its wave-length difference is found to be the same in value as that derived from the earlier theory, and as empirically found. Furthermore, in satisfactory agreement with experiment, Wentzel's theory yields the intensity distribution of the scattered radiation ; the *intensity ratio* between displaced and undis-

[5] A. H. Compton and A. W. Simon, *Phys. Rev.*, **25** (1924), p. 439.
[6] P. A. M. Dirac, *Proc. Roy. Soc., A* **111** (1926), p. 405.
[7] W. Gordon, *Zeitschr. f. Phys.*, **40** (1926), p. 117.
[8] G. Wentzel, *Zeitschr. f. Phys.*, **43** (1927), pp. 1 and 779.

placed lines,[9] which depends principally on the atomic number; and, quite generally, the dependence of the intensity on special properties of the scattering atoms.

## § 144. Dirac's Equations.

The manifold analogy between waves of light and of matter suggests an attempt at establishing a system of equations which should bear the same relation to the *wave equation of matter* as do *Maxwell's equations of the electromagnetic field* to the *wave equation of light*.[1] In point of fact, this attempt was successfully made in 1928 by *Dirac*,[2] starting, for that purpose, from the *relativistic generalization of the fundamental equation of wave-mechanics*.

In the theory of relativity, the three spatial coordinates are supplemented by a *fourth coordinate* which stands on just the same footing: this coordinate represents the *time* multiplied by the *velocity of light* and the imaginary quantity $\sqrt{-1}$, *i.e.*, by the product $i\,c\,t$. These four coordinates are correlated with four component velocities $v_x$, $v_y$, $v_z$, and $i\,c$ (as appears on differentiating $i\,c\,t$ with respect to the time). If, therefore, we distinguish *four component momenta*, which may be denoted by $p_1$, $p_2$, $p_3$, and $i\,p_0$, we have

$$(1) \qquad p_1 = m\,v_x,\ p_2 = m\,v_y,\ p_3 = m\,v_z,\ p_0 = m\,c\,.$$

Hence

$$-\,p_0^2 + p_1^2 + p_2^2 + p_3^2 = -\,m^2\,(c^2 - v^2)\,,$$

or, denoting the stationary mass, as defined in § 88 eqn. (6), by $m_0$,

$$(2) \qquad -\,p_0^2 + p_1^2 + p_2^2 + p_3^2 + m_0^2\,c^2 = 0\,.$$

Now it has been shown in an earlier section [§ 98 formula (6)] that the *momentum components* $p_1$, $p_2$, $p_3$ may be regarded as *operators*, *viz.*, as the partial derivatives with respect to the spatial coordinates, multiplied by $h/2\,\pi\,i$. If we interpret $i\,p_0$ similarly as the differential coefficient with respect to $i\,c\,t$, multiplied by $h/2\,\pi\,i$, we find

$$(3) \qquad p_0 = -\,\frac{h}{2\,\pi\,i\,c}\,\frac{\partial}{\partial t}\,, \quad p_1 = \frac{h}{2\,\pi\,i}\,\frac{\partial}{\partial x}\,, \text{ etc.}$$

[9] More accurately speaking, it is a question of the intensity ratio between the total displaced and the undisplaced radiation. From the point of view of the older quantum theory, the problem of the undisplaced lines was investigated principally by *G. E. M. Jauncey* [*Phys. Rev.*, **25** (1925), pp. 314 and 723].

[1] *Cf.* a paper by *E. Guth* in the *Anzeiger der Wien. Akad. der Wissen., math.-nat. Klasse*, 1928, p. 87.

[2] P. A. M. Dirac, *Proc. Roy. Soc., A* **117** (1928), p. 610, and **118** (1928), p. 351.

Let us put symbolically in this sense

$$(4) \qquad -p_0^2 + p_1^2 + p_2^2 + p_3^2 + m_0^2 c^2 \equiv g ;$$

then eqn. (2), as applied to Schrödinger's field-scalar,[3] gives the relation

$$(5) \qquad g S = 0 .$$

Thus, in the sense of eqns. (3), we obtain the *relativistic generalization of the fundamental equation of wave-mechanics*, which was first developed by *Schrödinger* [4] and *Gordon*,[5] in the form

$$(6) \qquad \nabla^2 S - \frac{1}{c^2}\frac{\partial^2 S}{\partial t^2} - \frac{4\pi^2 m_0^2 c^2}{h^2} S = 0 .$$

Dirac arrived now at his new equations by introducing an *operator* ($f$) which shall *depend linearly* on the *same* operators as those to which the operator $g$ stands in a *quadratic* relation. He made therefore the condition

$$(7) \qquad f \equiv p_0 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \alpha_4 m_0 c ,$$

and he postulated that the *product* of this operator with a second one ($f'$), differing from $f$ only in the sign of the first term,[6] should give the operator $g$. In the absence of a field, the relation

$$(8) \qquad f S = 0$$

would represent the new equation desired.

The postulate

$$(9) \qquad f f' = g$$

requires, on mathematical grounds, that the $\alpha$'s shall be matrices with four rows, and, consequently, the function $S$ must depend not only on $x$, $y$, $z$, $t$, but also on a *fifth variable* which, however, can assume only *four different values*. The dependence of the field-scalar $S$ on the fifth variable, which may be denoted by $u$, may be expressed by writing

$$(10) \qquad S(x, y, z, t, u) \equiv \psi_u(x, y, z, t) .$$

Thus, instead of one function depending on five variables, we have four functions which depend *only* on the spatial coordinates and the time.

It had previously been shown by Schrödinger and Gordon

[3] The operator $m_0^2 c^2$ denotes, of course, *multiplication* by $m_0^2 c^2$.

[4] E. Schrödinger, *Ann. der Phys.*, **81** (1926), p. 109; *vide* also "*Collected Papers on Wave Mechanics*."

[5] W. Gordon, *Zeitschr. f. Phys.*, **40** (1926), p. 117.

[6] Hence

$$f' \equiv -p_0 + \alpha_1 p_1 + \alpha_2 p_2 + \alpha_3 p_3 + \alpha_4 m_0 c .$$

that, in an *electromagnetic field*, the operators $p_0$, $p_1$, $p_2$, $p_3$ must be replaced as follows :—

$$(11) \qquad P_0 = p_0 + \frac{e}{c} A_0 ; \; P_1 = p_1 + \frac{e}{c} A_1 ; \; \textit{etc.} ;$$

where $A_0$ is the *electrostatic potential*, and $A_1$, $A_2$, $A_3$ the three components of the *vector potential*. As we know,[7] these quantities are connected with the components of the *electric field-strength* ($E$) and of the *magnetic field-strength* ($H$) by the relations

$$(12) \quad \begin{cases} E_x = -\dfrac{\partial A_0}{\partial x}, \qquad E_y = -\dfrac{\partial A_0}{\partial y}, \qquad E_z = -\dfrac{\partial A_0}{\partial z} ; \\ H_x = \dfrac{\partial A_3}{\partial y} - \dfrac{\partial A_2}{\partial z}, \; H_y = \dfrac{\partial A_1}{\partial z} - \dfrac{\partial A_3}{\partial x}, \; H_z = \dfrac{\partial A_2}{\partial x} - \dfrac{\partial A_1}{\partial y} . \end{cases}$$

Analogously to eqn. (7), Dirac defines an operator $F$ by the condition

$$(13) \qquad F \equiv P_0 + \alpha_1 P_1 + \alpha_2 P_2 + \alpha_3 P_3 + \alpha_4 m_0 c ,$$

where $\alpha_1$, $\alpha_2$, $\alpha_3$, $\alpha_4$ represent the *matrices* of four rows derived by him. The postulate that, in analogy to eqn. (8), the application of the operator $F$ to the field-scalar shall yield the value zero, or that

$$(14) \qquad F S = 0 ,$$

leads to *Dirac's equation* in the form [8]

$$(15) \quad \begin{cases} (P_0 + m_0 c)\, \psi_1 + (P_1 - i P_2)\, \psi_4 + P_3 \psi_3 = 0 , \\ (P_0 + m_0 c)\, \psi_2 + (P_1 + i P_2)\, \psi_3 - P_3 \psi_4 = 0 , \\ (P_0 - m_0 c)\, \psi_3 + (P_1 - i P_2)\, \psi_2 + P_3 \psi_1 = 0 , \\ (P_0 - m_0 c)\, \psi_4 + (P_1 + i P_2)\, \psi_1 - P_3 \psi_2 = 0 . \end{cases}$$

Bearing in mind the significance of the operators $P$, we can derive from these equations other equations, of the form [9]

[7] *Cf.* Vol. I., § 61.

[8] *Cf.* C. G. Darwin, *Proc. Roy. Soc.*, *A* **118** (1928), p. 654, eqns. (2.2).

[9] In deriving eqn. (16) from eqns. (15), we may follow Darwin by multiplying the first equation of the quadruplet by the operator $(-P_3)$, the second by the operator $(-P_1 + i P_2)$, and the third by the operator $(m_0 c - P_0)$, and finally by adding together. We must, of course, remember that the multiplication of operators is not commutative. For example, $P_1 P_2$ is certainly not equal to $P_2 P_1$, but we have

$$P_1 P_2 - P_2 P_1 = \frac{e}{c} \frac{h}{2\pi i} H_3 ,$$

and similarly,

$$P_0 P_1 - P_1 P_0 = \frac{e}{c} \frac{h}{2\pi i} E_1 .$$

$$(16)\qquad \begin{cases} [-P_0^2+P_1^2+P_2^2+P_3^2+m_0^2c^2]\,\psi_3 \\ \quad +\dfrac{e}{c}\dfrac{h}{2\pi}(H_1\psi_4 - iH_2\psi_4 + H_3\psi_3) \\ \quad +\dfrac{e}{c}\dfrac{h}{2\pi}i(E_1\psi_2 - iE_2\psi_2 + E_3\psi_1) = 0\,. \end{cases}$$

Similar equations may be derived for the other $\psi$-functions.

Now the expression in square brackets, which may be denoted by $G$, corresponds exactly to the previously used operator $g$, except that here the operators $p$ are replaced by the operators $P$. Schrödinger and Gordon had, in fact, already developed the wave-mechanical fundamental equation in an *electromagnetic field* in the form

$$(17)\qquad GS = 0\,.$$

As eqn. (16) shows, Dirac obtains essentially the same formula for the behaviour of a fundamental particle of matter (electron or proton) in an electromagnetic field, but with the addition of *two accessory terms*. A cursory inspection of these shows us that an *electron* in a magnetic field behaves as a *small magnet*, so that Dirac's equations *of themselves* account for *electron spin*, without the necessity of formulating a separate hypothesis for this. For the rest, it may be observed that these equations also indicate an *imaginary electric moment* [10]; whether any physical significance attaches to this unreal moment must remain as yet undetermined.

It has been a great achievement for Dirac's theory that, with its aid, *Darwin* [11] and *Gordon* [12] have effected an absolutely exact derivation of *Sommerfeld's* formulæ for the *fine structure of the hydrogen spectrum* (*cf.* § 104). Previous attempts, based on Uhlenbeck and Goudsmit's hypothesis, invariably led only to a first approximation at these formulæ.[13]

### § 145. The Transformation of Matter into Light.

The emission of light by atoms shows that energy inherent in *matter* can be transformed into the *energy of electromagnetic waves*. Moreover, as has been frequently

[10] This and the magnetic moment are found to have the same magnitude of $eh/4\pi m_0 c$, or one magneton (*cf.* § 116).

[11] *L.c.*

[12] W. Gordon, *Zeitschr. f. Phys.*, **48** (1928), p. 11.

[13] It is important to note that Dirac's equations, unlike Schrödinger's scalar equation, are applicable only to fundamental particles of matter, *i.e.* only to electrons and protons, but not to composite particles such as, for example, the $\alpha$-particle which (*cf.* § 133) has no resultant spin.

stated already, the theory of relativity leads to the important conclusion that to all matter as such is to be ascribed a so-called *proper energy*, which is determined by the enormous product of the *mass* and the *square of the velocity of light*. Within every milligram of matter slumbers the colossal energy of 22·5 million Kcal. Consequently, there arises the significant question whether not only the intra-atomic energy, but also the proper energy as such, can be transformed into light (in the widest sense of the term). In other words, is it possible to effect an *annihilation* of matter with a simultaneous generation of electromagnetic wave radiation? In point of fact, astronomical considerations have answered this question in the affirmative, and have taught that the *dissolution of matter into light* represents, as it were, the *fundamental phenomenon of cosmos*.

We know from experiments on radiation that the sun emits $4{\cdot}0 \times 10^{33}$ ergs in every second.[1] Since the mass of the sun is $2{\cdot}0 \times 10^{33}$ grams, the *loss in energy* referred to *one gram of solar matter* amounts approximately to *two ergs in a second*, or to $6{\cdot}3 \times 10^{7}$ ergs in a year. It is also known from astrophysical observations that the loss of energy by stars is much more turbulent in the early stages of their development, and may amount to as much as 700 ergs per gram and second. On the other hand, with stars that are still older than our sun (already of a ripe old age), the loss of energy is slower; *e.g.*, the star Krüger 60 radiates only 0·08 ergs per gram and second.

Now a gram of the best charcoal, when burnt, produces 8,000 cal., or $3{\cdot}3 \times 10^{11}$ ergs. The transformation of a gram of *hydrogen* into *helium* (*cf.* § 123) would generate $6{\cdot}8 \times 10^{18}$ ergs. Again, the dissolution of a gram of *matter*, *i.e.*, its transformation into radiation, would liberate $9 \times 10^{20}$ ergs. Therefore, for a supply of 2 ergs per gram and second the energy produced in the combustion of the charcoal would suffice for about 5,000 years; in the helium synthesis, for $1{\cdot}1 \times 10^{11}$ years; and in the dissolution of matter for about $15 \times 10^{12}$ years.

On the other hand, three different astronomical methods are available for determining with fair accuracy the *age of the sun*, or of the stars in our universe,[2] and calculations

[1] The distance of the earth from the sun is about $1{\cdot}5 \times 10^{13}$ cm., so that a sphere described about the sun with this radius has a surface of $2{\cdot}83 \times 10^{27}$ cm.$^2$ A square centimetre of the earth's surface receives 2·0 cal. in a minute by direct solar radiation, and thus a sixtieth of that amount in a second.

[2] With reference to this, and to this section in general, *vide* J. H. Jeans, *Nature*, **121** (1928), p. 462, and **122** (1928), p. 689.

effected according to the three methods *agree* in fixing the *age as between five* and *ten billion* ($10^{12}$) *years*. In passing, it may be noted that the first of these methods is based on the *progressive deformation*, produced in the originally circular *orbits of double stars*, by the continual perturbations for which the passage of other stars through their neighbourhood is responsible. The second method depends on the progressive *displacement* of the *lighter* with respect to the *heavier stars*, which is caused by the action of external forces on a *star cluster*. Finally, the third method utilizes the fact that the stellar system, like a gas, must ever tend to a state of "*equipartition*," the kinetic energies of the heavier and of the lighter stars continually approximating to one another.[3] The deformations of the orbits of double stars observed at the present time; the displacements already detected in star clusters (*e.g.*, within the star cluster in the Great Bear); and the approximation at equipartition hitherto attained; all these data can be used to derive the ages of the stars, with the result quoted above.

In earlier times, the solar radiation (referred to the gram) must have been more intensive than at the present day, but, even if we disregard this fact, the *helium synthesis* proves to be an entirely *insufficient* source of energy (by at least fifty times) when we bear in the mind the sun's age as calculated. Furthermore, apart from the dissolution of matter, it is difficult to conceive of an energy source more productive than the helium synthesis. Considerations on the packing effect (§ 123) demonstrate clearly that the energy, referred to a gram, generated in the synthesis of higher atomic nuclei from helium nuclei would be relatively small, as compared with the energy created in the synthesis of helium itself. A final decisive factor is that, consequent on the very *special position of hydrogen*,[4] *Eddington's* theory of the structure of the stars [5] leads necessarily to the conclusion that, originally, the sun could not have contained more than 10% of hydrogen. This means that the productiveness of a conjectural helium synthesis

[3] The process of equalization consists in the quickening up of the motions of the smaller, and the slowing down of those of the larger stars.

[4] Whereas with all the other elements the nuclear charge number is never more than half the atomic weight, with hydrogen both numbers are equal. As regards the special position of hydrogen, *cf.* the discussion in the article by E. Freundlich in "*Ergebnisse der exakten Naturwissenschaften*" (Berlin; Springer, 1927), Vol. VI.

[5] *Vide* A. S. Eddington, "*The Internal Constitution of the Stars*" (Cambridge; University Press, 1926).

would be reduced to 0·2% of the actual requirements of the sun.

On the other hand, if we look for the source of solar energy in the *dissolution of solar matter*,[6] at the *present* rate of evolving energy of 2 ergs per gram and second, this source would suffice (as has been stated) for $15 \times 10^{12}$ years, *i.e.*, for a period amounting to twice or thrice the age of the sun. Due account is thus paid to the circumstance that, in earlier times, the solar radiation referred to the gram was more intensive.

The assumption of a continuous *annihilation of stellar matter* [7] is supported by the fact that some 90% of the *young stars* possess *masses* lying between two and a half and five and a half times the mass of the sun, whereas the *average* mass of *all* the stars is certainly *smaller* than the solar mass. It seems, therefore, incontrovertible that, in the course of their development, the stars lose the greater part of their mass.

According to the general theory of relativity, the universe would occupy a finite space which permits of no distances larger than about $10^9$ light-years, with a total number of approximately $10^{24}$ stars, and a total mass of some $10^{57}$ grams.[8] Similarly, the history of the universe as such may perhaps be limited to a period which exceeds the time of $10^{13}$ years, previously quoted, by (at most) a few orders of magnitude, and within which the *dissolution of the universe* is inevitably perfecting itself.

[6] The suggestion of an annihilation of matter was first put forward by *Jeans*, and subsequently developed by *Eddington*. Obviously we must regard annihilation as the coalescence of a single proton and a single electron, with a resultant mutual destruction of charge and hence also of mass. The converse possibility of a *reconstruction of matter from starlight*, in consequence of successive *increases in frequency of light-quanta* due to collisions with extremely rapidly moving particles of matter, has been discussed by the author of this book. *Vide* A. Haas, *Sitz.-Ber. der Akad. Wiss. Wien, math.- naturwiss. Kl., Abt. IIa*, **135** (1926), p. 647 ; *Scientific Monthly*, **26** (1928), p. 140.

[7] Astronomical computations (*cf.* Jeans, *l.c.*) show that, were matter to be completely dissolved, the temperature of the universe would be thereby raised to about 10° above the absolute zero.

[8] These numbers result from estimates made by the American astronomer, *Hubble*, on the basis of Einstein's theory. *Cf.* E. Hubble, *Astrophys. Journ.* **64** (1926), p. 369.

# PART IV

# THEORY OF HEAT

## Chapter XVIII

# STATISTICS

### § 146. Statistical Probability

By resolving the object of physical processes, so-called *matter*, into an *assembly of individual particles*, it becomes possible for theoretical physics to investigate all physical phenomena and states by means of methods which are known quite apart from the province of physics as *statistical methods*. Among the manifold kinds of problems known to general statistics, problems of the following nature are of particular importance to physics. Let us imagine an assembly of individuals to be given, and let us consider one definite property of the individuals ; [1] we suppose that this property can be described in terms of one or more definite quantities which can be *graduated in intervals* in some way or another. In this way we obtain a simple or manifold plurality of states to be distinguished from each other. The question then arises as to how, under any given conditions, the *individuals* will be *distributed among the different states*.

Thus, for instance, the children born in a year in a large town represent an assembly of individuals. One of their many properties which we may consider is their weight at birth. This property can be described by a single quantity, the weight, which may be graduated in various ways, *viz.*, in tenths of kilograms or in decagrams, *etc.* A definite state is then represented, for example, by a weight of between 3·2 and 3·3 kg. Another property of the children, for the description of which, however, two quantities are necessary, is the age of their parents at the time of birth. In this case we have to deal with a double plurality of states, the age both of the father and of the mother being graduated, for instance, in years or months. A definite state would then be that in which the father's age lies in the interval between thirty-three and thirty-four years, and the mother's age

[1] *I.e.*, one among many ; *e.g.*, in human beings the height of the body, or the age, or the financial status, *etc.*

between twenty-one and twenty-two years. A further example of a distribution of individuals among states is represented by the distribution of the population of a country according to their yearly income. The distribution is then bound by the condition that, if we multiply the number of people belonging to each interval of income by the value of the income for that interval, and then take the sum over all the intervals, the result must give the total annual income of the population.[2]

Just as general in character as the problem of individual distribution are some concepts to which the treatment of this problem leads. Let us consider an assembly of $N$ individuals, and suppose the different intervals of state to be distinguished by indices ; thus $N_i$ will be the number of individuals in the state characterized by the index $i$ for a certain condition of distribution of states. A definite *distribution of states* is then described by the numbers $N_1, N_2, \ldots N_i, \ldots$, *etc.*, with the qualification, however, that the sum of all the numbers $N_i$ over all the states must give the total number of all the individuals. We must therefore have

$$\sum N_i = N. \tag{1}$$

Should we include states which contain no individuals at all,[3] this naturally would not in the least affect the consideration, because for such state the corresponding number $N_i$ vanishes.

A distribution of states is therefore completely determined as soon as we know *how many* individuals the different states contain. *Which* the individuals are is a matter of indifference. Hence, if two individuals in different states are interchanged, *no new* distribution arises, for the numbers $N_1, N_2, \ldots N_i, \ldots$, *etc.*, are not altered. A new "*complexion*" does, however, arise ; and thus, in general, a definite distribution of states is associated with many complexions. The number of complexions is termed the *statistical probability* of the distribution of states in question.

The statistical probability, which must always be a positive integer, is obviously given by the number of *permutations* of $N$ individuals taken altogether, omitting, however, those arising from an interchange of two individuals in the same state. Now the total number of all possible permutations

[2] Naturally not with absolute accuracy, for the intervals have been chosen as finite.

[3] This would be the case, for instance, if we took into consideration weights below 1 kg. or above 8 kg. in discussing the weights of newly born infants.

of $x$ elements is $x!$ ("factorial $x$"), where, by definition of the symbol,

$$x! = 1 \cdot 2 \cdot 3 \ldots (x-2) \cdot (x-1) \cdot x. \tag{2}$$

The total number of permutations of $N$ individuals taken altogether is therefore $N!$. The number of permutations resulting from an interchange of the individuals in the first state is $N_1!$, *etc.* Hence, we obtain the statistical probability of a given distribution of states on dividing $N!$ in the first place by $N_1!$, because all the combinations arising from interchanges between individuals in the first state represent, as a whole, only one complexion (and not $N_1!$). Then we have to divide by $N_2!$, and so on. If $W$ be the statistical probability, we have therefore [4]

$$W = \frac{N!}{N_1!\ N_2! \ldots N_i! \ldots}. \tag{3}$$

We shall now introduce new quantities, defined by the equation

$$w_i = \frac{N_i}{N}; \tag{4}$$

*i.e.*, the quantity $w_i$ gives the fraction of the individuals occurring in the state characterized by the index $i$. According to eqn. (1), we must have

$$\sum w_i = 1. \tag{5}$$

The expression for the statistical probability can be considerably simplified with the help of an easily derivable mathematical *formula of approximation*. This formula concerns the logarithm of $n!$. We suppose the curve

$$y = \log_e x$$

to be constructed in an $x$-$y$-coordinate system, and at all points of which the $x$-coordinate is *integral*, but not greater than a whole number $n$, the corresponding ordinates to be drawn. We obtain then (Fig. 123) a figure $A\ B\ C$, the coordinates of the point $A$ being $x = 1$, $y = 0$; those of the

[4] *E.g.*, in an assembly of four individuals, $a\ b\ c\ c$, the two $c$'s form a subgroup of two individuals, because an interchange between them does not lead to any new complexion. The possible complexions are

| | | |
|---|---|---|
| $a\ b\ c\ c$ | $a\ c\ b\ c$ | $a\ c\ c\ b$ |
| $b\ a\ c\ c$ | $b\ c\ a\ c$ | $b\ c\ c\ a$ |
| $c\ a\ b\ c$ | $c\ a\ c\ b$ | $c\ b\ a\ c$ |
| $c\ b\ c\ a$ | $c\ c\ a\ b$ | $c\ c\ b\ a$ |

The number of these complexions is, in fact, equal to $4!/2!$, or 12.

point $B$, $x = n$, $y = 0$; and those of the point $C$, $x = n$, $y = \log_e n$. The area of this figure is given by

$$F = \int_1^n \log_e x \, dx .$$

We carry out the integration, in the usual way, with the help of the identity [5]

(6) $$\log_e x \, dx = d(x \log_e x) - dx ,$$

and find

(7) $$F = n \log n - n + 1 .$$

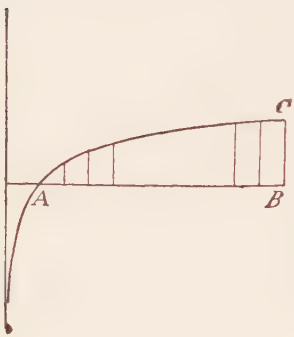



Fig. 123.

On the other hand, the area of the figure $A\ B\ C$ is divided into a number of *strips* by the ordinates from the $n$ points with whole number abscissæ. Apart from the lower values of $x$, these strips can be regarded as small *rectangles*, for then the logarithmic curve runs approximately parallel to the $x$-axis. The width of these rectangles is always unity, and their height $\log_e x$ ($x = 2$ to $n$). On this assumption, the area of the figure appears as the sum

(8) $$\log_e 2 + \log_e 3 + \ldots + \log_e (n-1) + \log_e n = \log_e [1 \,.\, 2 \,.\, 3 \ldots (n-1) \,.\, n] = \log_e (n\,!).$$

If $n$ be large, the error due to our regarding the strips as rectangles even for small values of $x$ vanishes. Hence, for large values of $n$, expressions (7) and (8) may be taken as identical. Since the number 1 may then be disregarded in comparison with $n$, for large values of $n$ we find the approximate formula [6]

[5] We have

$$dx = x \frac{1}{x} \, dx = x \, d(\log_e x) .$$

[6] This formula represents a special case of a more exact approximation formula, known as *Stirling's formula* (*vide* textbooks on probability). According to Stirling's formula,

(9b) $$\log_e (n\,!) = n \log_e n - n + \frac{1}{2} \log_e (2 n \pi) .$$

For large values of $n$, this formula is extraordinarily accurate; even for $n = 10$ the error is less than 1%. Eqn. (9) is, of course, less accurate than Stirling's formula, the last term in the latter being omitted. For large values of $n$, however, it is easy to see that this term is quite unimportant.

$$\log_e (n\,!) = n \log_e n - n\,. \tag{9}$$

This approximation formula may be applied to eqn. (3), if we assume that the large majority of individuals occur in states each of which contains a large number of them. For, even if there be only a few individuals in any state $k$, but many individuals in a state $h$, $\log_e (N_k\,!)$ may simply be neglected as compared with $\log_e (N_h\,!)$, inasmuch as, with increasing $x$, $x\,!$ increases more rapidly than $e^x$.[7] Applying, therefore, the approximation formula to eqn. (3), we find

$$\log_e W = N \log_e N - N - \sum N_i \log_e N_i + \sum N_i\,,$$

or, bearing in mind eqns. (1) and (4),

$$\begin{aligned}\log_e W &= N \log_e N - \sum N w_i (\log_e N + \log_e w_i) \\ &= N \log_e N - N \log_e N \sum w_i - N \sum w_i \log_e w_i\,.\end{aligned}$$

Finally, taking eqn. (5) into account, we obtain the simple relation

$$\log_e W = - N \sum w_i \log_e w_i\,. \tag{10}$$

This formula determines the statistical probability for a given distribution of states.

Now, the larger the statistical probability of any distribution, the larger also is the probability that any given distribution will change into it. If $A$, $B$, $C$ are three distributions, and their statistical probabilities are $W_1$, $W_2$, $W_3$, where $W_2 < W_1 < W_3$, the transition from $A$ to $C$ must be more probable than that from $A$ to $B$. Since, however, these distributions are quite arbitrary, the probability is that, on a change of distribution, the *statistical probability* of a system will *increase*.

On the other hand, it is evident that there must be an *upper limit* to the statistical probability of a system, for, by definition, this cannot possibly be larger than the *total number of all complexions*. This number is given by the $N^{\text{th}}$ power of $z$, when $N$ individuals are distributed over $z$ states.[8]

[7] We have

$$\frac{x\,!}{e^x} = \frac{1}{e} \cdot \frac{2}{e} \cdot \frac{3}{e} \cdots \frac{x-1}{e} \cdot \frac{x}{e}.$$

From the third term onwards, however, all the factors are larger than unity and increase throughout.

[8] For, if there were only two individuals, there would be $z$ possible states for the first of them. To each single possible state, however, there correspond $z$ further complexions, because, even when the state of the first individual is fixed, there are still $z$ possible states for the second. With two individuals, therefore, the total number of all complexions is given by $z^2$; and hence, in general, with $N$ individuals by $z^N$.

The number of *possible* complexions is, of course, smaller when the distribution is subject to conditions which can be satisfied by only some of the complexions. In every case, however, there is an upper limit, and this limit determines the *most probable distribution*, towards which the system *apparently strives*, and which, on account of its particularly *large number of complexions, occurs most frequently*.[9]

## § 147. The Modulus of Distribution.

The most important condition governing a distribution of individuals over different states is as follows : Let the system possess a certain *amount* $B$ of any quantity, this amount being equal to the *sum of the component amounts* associated with the separate individuals. Let the component amount associated with the individual depend, on the one hand, on the state in which the individual occurs, and on the other hand, on quantities which, at any moment, are the same for all individuals in the system, and may be termed the *parameters of the system*. Denoting these parameters, $r$ in number, by $u_1, u_2, \ldots u_r$, and by $\beta_i$ the component amount associated with an individual in the definite state characterized by the index $i$, then $\beta_i$ is a function only of the parameters, so that

$$\beta_i = f(u_1, u_2, \ldots u_r). \tag{1}$$

Again, if $N_i$ be the number of individuals in the state characterized by the index $i$, we must have

$$B = \sum N_i \beta_i. \tag{2}$$

We meet, now, with a *special* case of great importance when the distribution is such that its *statistical probability* represents an *extreme value* for those *variations of the distribution* in which the total amount of the quantity and the number of individuals remain constant. A distribution which

[9] The correctness of this theorem is, indeed, confirmed by experiences of everyday life. If we place layers of blue and yellow powder in a glass tube, the contents of the tube after a little shaking will appear uniformly green. But, however much we shake, we shall never be able to make the lower layer once more appear blue and the upper yellow. Order changes to disorder by itself, but the reverse process, the fortuitous appearance of order, is only to be observed with a very small number of individuals. The chance that four men who die successively in a town depart from this life in the alphabetical order of their names is, on an average, 1 in 24 ($24 = 4!$). In a series of ten deaths, the probability of such an alphabetical arrangement is less than a millionth, in a series of twenty deaths less than 1 in $10^{18}$, and in a series of thirty deaths less than 1 in $10^{30}$.

satisfies both this main condition and the subordinate conditions is termed a *canonical distribution*. Since it is more convenient to calculate with the logarithm of the probability rather than with the probability itself, the main condition may be expressed by the equation

$$\delta\,(\log_e W) = 0\,; \tag{3}$$

and the subordinate conditions by

$$\delta N = 0, \qquad \delta B = 0\,. \tag{4}$$

In accordance with § 146 eqns. (4) and (5), the constancy of the number of individuals may be expressed by the formula

$$\sum \delta w_i = 0\,. \tag{5}$$

Since the variation refers only to the quantities $w_i$, we have

$$\delta B = \sum \beta_i\, \delta N_i\,. \tag{6}$$

On account of the constancy of $N$, the subordinate condition of the vanishing of the variation of $B$ may be expressed by the equation

$$\sum \beta_i\, \delta w_i = 0\,. \tag{7}$$

It follows from § 146 eqn. (10) that, since $N$ is constant, the variation of the probability logarithm is given by

$$\delta\,(\log_e W) = -N \sum \log_e w_i\, \delta w_i - N \sum \delta w_i\,. \tag{8}$$

Since the last sum vanishes [by eqn. (5)], the principal condition of a canonical distribution is expressed by the equation

$$\sum \log_e w_i\, \delta w_i = 0\,. \tag{9}$$

Now, in accordance with the rules of the calculus of variations, we can combine this equation, expressing the principal condition, with eqns. (7) and (5), expressing the subordinate conditions, to form a single equation. This we do by multiplying each of eqns. (7) and (5) by a factor (at first left undetermined), and then adding them to eqn. (9). (The process is the same as in Lagrange's method of undetermined multipliers.[1]) Denoting the two undetermined factors by $\lambda_1$ and $\lambda_2$, we obtain thus

$$\sum (\log_e w_i + \lambda_1 \beta_i + \lambda_2)\, \delta w_i = 0.$$

As the variations are quite arbitrary, this equation can only be satisfied when the expression in brackets vanishes.[2] We must therefore have

[1] *Vide* § 20.
[2] *Cf.* the analogous considerations in § 20.

$$\log_e w_i = -\lambda_2 - \lambda_1 \beta_i ,$$

or,

$$w_i = e^{-\lambda_2 + \beta_i \lambda_1} , \tag{10}$$

We shall now introduce two new quantities, $\psi$ and $\Theta$, defined by the relations

$$\lambda_1 = \frac{1}{\Theta}, \qquad \lambda_2 = -\frac{\psi}{\Theta}. \tag{11}$$

Then

$$w_i = e^{\frac{\psi - \beta_i}{\Theta}} . \tag{12}$$

The two quantities $\psi$ and $\Theta$ are determined in accordance with the subordinate conditions in such a way that

$$N \sum w_i \beta_i = B \tag{13}$$

and

$$\sum w_i = 1 . \tag{14}$$

From eqns. (14) and (12) we obtain the relation

$$e^{-\frac{\psi}{\Theta}} = \sum e^{-\frac{\beta_i}{\Theta}} , \tag{15}$$

while we find from eqns. (13) and (12) that

$$B = N \sum \beta_i \, . \, e^{\frac{\psi - \beta_i}{\Theta}} = \frac{N \sum \beta_i e^{-\frac{\beta_i}{\Theta}}}{e^{-\frac{\psi}{\Theta}}} ,$$

or, by eqn. (15),

$$\frac{\sum \beta_i e^{-\frac{\beta_i}{\Theta}}}{\sum e^{-\frac{\beta_i}{\Theta}}} = \frac{B}{N} . \tag{16}$$

This equation determines the quantity $\Theta$, which is known as the *modulus of distribution*. Once that this is given, $\psi$ can be calculated from eqn. (15). Introducing, for the simplification of subsequent calculations,

$$N \psi = \Psi , \tag{17}$$

we find by eqn. (15)

$$\Psi = -N \Theta \log_e \sum e^{-\frac{\beta_i}{\Theta}} . \tag{18}$$

The value of the probability logarithm for the case of canonical distribution follows from eqn. (12), and from § 146 eqn. (10):—

$$\log_e W = -N \sum \frac{\psi - \beta_i}{\Theta} e^{\frac{\psi - \beta_i}{\Theta}},$$

or,

$$\log_e W = -\frac{N\psi}{\Theta} \sum e^{\frac{\psi - \beta_i}{\Theta}} + \frac{N}{\Theta} \sum \beta_i e^{\frac{\psi - \beta_i}{\Theta}}.$$

According to eqns. (12) and (14), the first sum is equal to unity, and the second sum, by eqns. (12) and (13), is equal to $B/N$. Hence, it follows from eqn. (17) that, for a canonical distribution,

$$\log_e W = \frac{B - \Psi}{\Theta}. \tag{19}$$

Since $\beta_i$ depends only on the parameters,[3] we see from eqn. (18) that $\Psi$, and hence also $\psi$, can be a *function only of the parameters and the modulus of distribution*. In accordance with eqn. (16), this holds also for the quantity $B/N$ (the average value of the component amount), and, by eqn. (19), for the probability logarithm.

Now, in virtue of its definition, a canonical distribution is characterized by the constancy of the probability logarithm for small variations of the distribution. Obviously, therefore, an *especially large number of complexions* will have a value of the probability logarithm corresponding to a canonical distribution. Since, however, the number of complexions by means of which a distribution can be realized is simply the statistical probability, the extreme value of the *logarithm of the probability* for a canonical distribution must be a *maximum*.

A canonical distribution thus represents at the same time the most probable and the most frequent distribution. Moreover, it is distinguished by definition by the maintenance of its canonical character for infinitesimal variations. Otherwise, in accordance with what we have said above, variations in distribution are accompanied by a movement in the direction of increasing statistical probability rather than in the reverse direction. *Infinitesimal canonical variations*, on the other hand, occur without any change in the

[3] *Vide* eqn. (1).

statistical probability, and thus appear to be equally probable in either directions ; they are apparently *reversible*.[4]

## § 148. The Canonical Distribution of Energy.

In the following discussion, we shall apply our previous statistical considerations by identifying the quantity $B$ with *energy*, for which the *Law of Conservation* holds good. The canonical distribution of energy is therefore determined by the formula

$$w_i = e^{\frac{\psi - \epsilon_i}{\Theta}}, \tag{1}$$

where $\psi$ and $\Theta$ must satisfy the two conditions

$$\sum e^{\frac{\psi - \epsilon_i}{\Theta}} = 1 \tag{2}$$

and

$$N \sum \epsilon_i e^{\frac{\psi - \epsilon_i}{\Theta}} = E. \tag{3}$$

If, now, the parameters $u_1$, $u_2$, *etc.*, are varied by the infinitesimal amounts $du_1$, $du_2$, *etc.*, and if, as we have assumed, the quantity $\epsilon_i$ depends only on the parameters, then the variation in the energy of a single individual is given by

$$\frac{\partial \epsilon_i}{\partial u_1} du_1 + \frac{\partial \epsilon_i}{\partial u_2} du_2 + \dots .$$

For an infinitesimal variation of the parameters, the total variation in energy of a system in a state of canonical distribution is given by a summation over all the individuals. Hence, the total variation of the energy is represented by the expression

$$\sum \left\{ N_i \frac{\partial \epsilon_i}{\partial u_1} du_1 + N_i \frac{\partial \epsilon_i}{\partial u_2} du_2 + \dots \right\}.$$

In accordance with the axiomatic *Principle of the Conservation of Energy*, this quantity must be equal and of opposite sign to the *work* performed by the system as a result of the variation of the parameters. We obtain thus the relations

$$dA = - N\, du_1 \sum w_i \frac{\partial \epsilon_i}{\partial u_1} - N\, du_2 \sum w_i \frac{\partial \epsilon_i}{\partial u_2} - \dots . \tag{4}$$

[4] The variations must be regarded as being infinitesimal ; *cf.* the analogous considerations on the Principle of Virtual Displacements (§ 18).—For an exact proof of the assertion that the canonical distribution is also the most probable, *cf.* in particular the article by P. and T. Ehrenfest in the *Mathematische Enzyklopädie* IV., 2, II., part 6.

The summands in this equation may, however, be greatly simplified with the help of eqn. (2). In doing this, we must remember that $\epsilon_i$ is a function only of the $u$ values, whilst $\psi$ is a function only of the $u$ values and $\Theta$. We may therefore regard the summand expressions in eqn. (2) as functions of the $u$ values and $\Theta$, and hence differentiate them partially with respect to the parameters. In this way we obtain

$$\frac{1}{\Theta}\sum \frac{\partial(\psi - \epsilon_i)}{\partial u_1} e^{\frac{\psi-\epsilon_i}{\Theta}} = 0,$$

or,

$$\sum \frac{\partial \epsilon_i}{\partial u_1} e^{\frac{\psi-\epsilon_i}{\Theta}} = \frac{\partial \psi}{\partial u_1} \sum e^{\frac{\psi-\epsilon_i}{\Theta}}, \tag{5}$$

or, taking eqns. (1) and (2) into account,

$$\sum \frac{\partial \epsilon_i}{\partial u_1} w_i = \frac{\partial \psi}{\partial u_1}. \tag{6}$$

Eqn. (4) thus becomes, on replacing $N\psi$ by $\Psi$,

$$dA = -\frac{\partial \Psi}{\partial u_1} du_1 - \frac{\partial \Psi}{\partial u_2} du_2 - \ldots, \tag{7}$$

Since the quantity $\Psi$ thus plays the same part for the system as a whole as potential energy would, it is called the *free energy* of the system. [It follows, in fact, from § 147 eqn. (19) that $\Psi$ and $\Theta$ must always have the same dimensions as $B$, for $\log_e W$ must, of course, always be a pure number.]

We can write eqn. (7) in another important form, when we remember that, in accordance with what was said above, $\Psi$ must be determined by the parameters and the modulus of distribution. Thus

$$d\Psi = \frac{\partial \Psi}{\partial \Theta} d\Theta + \frac{\partial \Psi}{\partial u_1} du_1 + \frac{\partial \Psi}{\partial u_2} du_2 + \ldots \tag{8}$$

Hence, by eqn. (7),

$$dA = -d\Psi + \frac{\partial \Psi}{\partial \Theta} d\Theta. \tag{9}$$

But, by § 147 eqn. (18),

$$\Psi = -N\Theta \log_e \sum e^{-\frac{\epsilon_i}{\Theta}}, \tag{10}$$

and therefore

$$\frac{\partial \Psi}{\partial \Theta} = -N \log_e \sum e^{-\frac{\epsilon_i}{\Theta}} - \frac{N\Theta \sum \epsilon_i e^{-\frac{\epsilon_i}{\Theta}}}{\Theta^2 \sum e^{-\frac{\epsilon_i}{\Theta}}},$$

or, bearing in mind eqn. (10) and § 147 eqn. (16),

$$\frac{\partial \Psi}{\partial \Theta} = \frac{\Psi}{\Theta} - \frac{E}{\Theta}. \tag{11}$$

Hence, by eqn. (9),

$$dA = -d\Psi - \frac{E - \Psi}{\Theta} d\Theta. \tag{12}$$

If we suppose an *infinitesimally small quantity of energy* ($dQ$) to be introduced into the system considered,[1] then in general both the "*internal energy*" of the system will be thereby altered, and also *work* will be done by the variation of the parameters. According to the Principle of the *Conservation of Energy*, we must have

$$dQ = dE + dA. \tag{13}$$

In this expression $dE$ is a complete differential, because, by eqns. (3) and (8), $E$ is completely determined by the modulus of distribution and the parameters, and consequently depends only on the state of the system. On replacing $dA$ in eqn. (13) by its value as given in eqn. (12), we obtain the relation

$$dQ = d(E - \Psi) - \frac{E - \Psi}{\Theta} d\Theta. \tag{14}$$

We can easily see that the right-hand side of this equation becomes a *complete differential* when the equation is divided by $\Theta$. We have then

$$\frac{dQ}{\Theta} = \frac{d(E - \Psi)}{\Theta} - \frac{E - \Psi}{\Theta^2} d\Theta,$$

or

$$\frac{dQ}{\Theta} = d\left(\frac{E - \Psi}{\Theta}\right), \tag{15}$$

or, finally, taking § 147 eqn. (19) into account,

$$\frac{dQ}{\Theta} = d(\log_e W). \tag{16}$$

[1] $dQ$ is, of course, no more a complete differential than is $dA$.

From a mathematical point of view, the *modulus of distribution* thus represents an *integrating divisor*, by virtue of which the infinitesimal energy introduced becomes a complete differential. The function of which the total differential is the fraction $dQ/\Theta$, however, has (as we have shown in §§ 146 and 147) the property of appearing to alter in one sense only, and of remaining constant for reversible variations in distribution.

For a system the *energy* of which is *canonically distributed*, the following two important theorems therefore hold good :

I. If an infinitesimal quantity of energy be introduced into the system, while the latter performs work, the *difference between the energy introduced and the work performed* represents the *complete differential of a function* which depends only on the *instantaneous state of the system*, and is completely determined by the modulus of distribution and the parameters.

II. The *quotient of the infinitesimal quantity of energy introduced and the modulus of distribution* represents the *differential of a second function* which likewise depends only on the instantaneous state of the system, and is likewise completely determined by the modulus of distribution and the parameters. This function has the property of remaining *constant* in *reversible* processes, but otherwise it generally *increases*.

## § 149. Equilibrium and Modular Equalization.

When *several systems together* form a *separate whole*, this again may be regarded as a single system and associated with a definite *statistical probability*. The relation of the latter to the probabilities of the component systems is very simple. For the sake of simplicity, we shall first assume that we are dealing with only two component systems of given distributions, the statistical probabilities being $W_1$ and $W_2$. The given distribution of the first system can then be realized by $W_1$ complexions, and that of the second system by $W_2$ complexions. Since, however, each complexion of the first system can be combined with each of the $W_2$ complexions of the second, the total number of complexions through which the given distribution can be realized in the composite system made up of the two component systems is $W_1$ times $W_2$.[1]

[1] We might quite easily draw the following false conclusion. If the number of individuals in the two component systems be $N'$ and $N''$, and if the first state contain $N_1'$ and $N_1''$ individuals respectively, we might suppose the statistical

Since this consideration may be directly extended to any arbitrary number of component systems, the statistical probability of the whole is equal quite generally to the *product of the statistical probabilities of the component systems.* The logarithm of the probability of the composite system is therefore equal to the *sum of the probability logarithms of the component systems.* We have

$$\log_e W = \log_e W_1 + \log_e W_2 + \ldots \tag{1}$$

Like the component probabilities, the total probability must obviously have an upper limit which cannot be greater than the product of the upper limits of the component probabilities.[2] Hence, there is also a *maximum* for the probability logarithm of the composite system, and this will evidently be attained in a distribution, the probability logarithm of which does not vary when *infinitesimal energy interchanges* occur between the *component systems.*

If we consider two canonical component systems, which may be distinguished simply by the indices 1 and 2, and if a quantity of energy ($dQ$) be transferred from the first to the second system, then, by § 148 eqn. (16),

$$-\frac{dQ}{\Theta_1} + \frac{dQ}{\Theta_2} = d(\log_e W_1) + d(\log_e W_2)\,. \tag{2}$$

When

$$\Theta_1 = \Theta_2\,, \tag{3}$$

the right-hand side of the equation vanishes, and the sum of the probability logarithms of the two systems then remains constant. If, therefore, the composite system be made up of component systems, each of which is canonically distributed and has the *same value of the modulus of distribution*, the probability logarithm of the whole is not altered by infinitesimal interchanges of energy between the component systems. In that case, the component systems are said to be in a state of *statistical equilibrium.*

In statistical equilibrium, the probability logarithm represents a *maximum.* Since, under other circumstances, it seems to tend towards this, there is an apparent endeavour

probability of the composite system to be obtained by dividing $(N' + N'')!$, first by $(N_1' + N_2'')!$, then by $(N_2' + N_2'')!$, *etc.* This conclusion, however, would be quite wrong. For, if an individual in any state in the first component system be interchanged with an individual in the second, this gives rise not only to a new complexion, but also, in general, to a new distribution in the two component systems.

[2] Should certain conditions exist, it will, of course, be smaller than this product.

on the part of every composite system to produce statistical equilibrium between its component systems. When systems with different moduli of distribution are connected in such a way that a transfer of energy is possible between them, an *equalization in the values of the moduli of distribution* will, in general, occur.

Inasmuch as an equalization of this kind is always accompanied by an increase of the total probability logarithm, a change in state of a system is only *reversible* on two conditions. First, the distribution in the system itself must remain canonical; and secondly, during the change, any combinations with other systems into which the system enters must be associated with only *infinitesimal deviations from statistical equilibrium*. For the rest, in consequence of the existence of an equation of state (to be discussed later), a change in the modulus of distribution must entail a change in at least one of the other quantities of state. If, therefore, no change is apparent in the original conditions of two systems when they are brought in contact, it may be concluded that they both have the *same value* of the modulus of distribution.

Since, on an equalization of moduli, the total probability logarithm increases, it follows from eqn. (2) ($\Theta_1$ being not equal to $\Theta_2$) that

$$-\frac{dQ}{\Theta_1}+\frac{dQ}{\Theta_2}>0. \tag{4}$$

That is to say, that component system which *loses* energy must have the larger modulus. Rare as is a process in which the statistical probability diminishes, the *transfer of energy of its own accord from a system with a smaller to one with a larger modulus* must be just as *rare*.

If, in an equalization of moduli, the first system has a far larger number of individuals and a far greater energy than the second, the common value of the modulus after the equalization cannot differ appreciably from the original value ($\Theta^*$) of the modulus in the first system. In this special case, therefore, the process of equalization simply consists in an assumption by the second system of the modulus of distribution of the first. Compared with the second, the first system thus acts as a "*reservoir*" with a modulus value of $\Theta^*$.

### § 150. Homogeneous Systems.

As we have seen, the free energy of a system is completely determined by the modulus of distribution and the para-

meters, and the same must therefore hold for the partial derivatives of the free energy with respect to the parameters, *i.e.*, for the quantities

$$(1) \qquad U_1 = -\frac{\partial \Psi}{\partial u_1}, \qquad U_2 = -\frac{\partial \Psi}{\partial u_2}, \ldots$$

In accordance with § 148 eqn. (7), the dependency of the quantities $U_1, U_2$, *etc.*, defined by these equations, on the work performed for an infinitesimal variation of the parameters is given by the relation

$$(2) \qquad dA = U_1\, du_1 + U_2\, du_2 + \ldots$$

The quantities $U_1, U_2$, *etc.*, may therefore be termed generalized *statistical forces*, although their dimensions are by no means necessarily those of a force.

Since the statistical forces are thus completely determined by the modulus of distribution and the parameters, we have, if $r$ be the number of parameters, $r$ equations of the form

$$(3) \qquad \begin{cases} U_1 = f_1(\Theta, u_1, u_2, \ldots u_r), \\ U_2 = f_2(\Theta, u_1, u_2, \ldots u_r), \text{ etc.} \end{cases}$$

The number $r$ may be termed the number of *statistical degrees of freedom*. The quantities $\Theta$, $U_1$, $U_2$, $\ldots U_r$, $u_1$, $u_2$, $\ldots u_r$ are called the *quantities of state of the system*, and processes in which one or more of these quantities change are defined as *changes of state*. Owing to the existence of the $r$ "*equations of state*" (3), which can be solved with respect to the parameters, only $(r + 1)$ out of the $(2r + 1)$ quantities of state are mutually independent. The state of a system with $r$ degrees of freedom can therefore always be represented by $(r + 1)$ *independent variables of state*, to be chosen at will from among the quantities of state.

An important special case is presented by a system which has only *one degree of freedom*, the *single parameter* being defined by the *quotient of the volume* occupied by the system *and the number of individuals*.[1] A system of this nature is said to be *statistically homogeneous*. We have then for the single parameter

$$(4) \qquad u = \frac{V}{N}.$$

[1] $V$ denotes the space within which the individuals are contained: it must not be confused with the sum of the volumes occupied by the separate individuals.

When the parameter $u$ is increased, the system performs work against the *pressure* ($p$) acting upon its surface. This work may readily be calculated.

Let us consider an element ($df$) of the surface of the system. A force of magnitude $p \, . \, df$ will thus act upon the element. Let us suppose that, for an infinitesimal variation of the volume, the surface-element is displaced from the position represented by the interval $AB$ to that indicated by the interval $A'B'$ (Fig. 124), the mid-point of the element being thereby moved from $O$ to $O'$. The *work* performed in the displacement of the surface-element is then equal to $p \times df \times OO' \times \cos \gamma$, where $\gamma$ is the angle included between the interval $OO'$ and the normal to the surface. (For an infinitesimal displacement, the normals to the surfaces $AB$ and $A'B'$ will evidently differ in direction by only an infinitesimal amount).

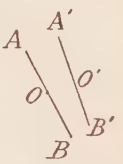


FIG. 124.

But $df \times OO' \times \cos \gamma$ is the volume of the infinitesimally small cylinder of which the ends are formed by the two equally large surface-elements $AB$ and $A'B'$, and the height is equal to $OO' \times \cos \gamma$. The work performed in the displacement of the surface-element is thus equal to the product of the pressure and the volume of the cylinder formed in the displacement. If we take the sum of this volume over the whole surface of the system, we obtain the total increase in volume, which may be denoted by $dV$. Hence, the total work ($dA$) performed by the system for an increase in its volume by the amount $dV$ is given by

$$dA = p \, dV = N p \, du \, . \tag{5}$$

By eqn. (2), therefore,

$$U = N p \, , \tag{6}$$

and hence, according to eqn. (1),

$$p = -\frac{1}{N} \frac{\partial \Psi}{\partial u} \, , \tag{7}$$

or, by eqn. (4),

$$p = -\frac{\partial \Psi}{\partial V} \, . \tag{8}$$

When the number of individuals is constant, the volume is proportional to the parameter $u$, and therefore, in a homogeneous system, instead of the parameter the *volume* itself can be regarded as a *quantity of state*. In a canonical distribution, pressure, volume, and modulus of distribution

are thus always connected by an *equation of state*. It follows also that none of these three quantities can vary without at least one of the other two varying. Hence if, when two homogeneous systems are brought into contact in such a way that a transfer of energy is possible, neither volume nor pressure alter, we must conclude that both systems have the *same* modulus of distribution.

## § 151. Cyclic Processes.

As has been shown in § 148, the application of the principle of energy to canonical systems leads to the two principal equations

$$dQ = dE + dA\,, \tag{1}$$

and

$$\frac{dQ}{\Theta} = d(\log_e W)\,. \tag{2}$$

We can at once see that, in two special cases, the equations must assume an especially simple form.

In the first place, if the reversible change of state of the system occurs without any transfer of energy to or from outside sources, *i.e.*, if it occurs *adiabatically* (because no energy then passes through the surface of the system [1]), $dQ$ is to be equated to zero in the integration of eqn. (1). Since, however, the internal energy depends only on the instantaneous state, integration of eqn. (1) in this case gives

$$\sum dA = E_1 - E_2\,, \tag{3}$$

where $E_1$ and $E_2$ are the values of the internal energy in the initial and the final states.

Secondly, if the reversible change of state occurs in such a way that the value of the modulus of distribution remains constant, eqn. (2) then assumes a simple form. For, if we denote the statistical probabilities of the initial and the final states by $W_1$ and $W_2$, it follows from eqn. (2) that, since $\Theta$ is constant,

$$\sum dQ = \Theta \log_e \frac{W_2}{W_1}\cdot \tag{4}$$

Like the quantity $E$, the quantity $W$ (according to § 148) also depends solely on the instantaneous state of the system. For the moment, we shall define a reversible process in which the modulus of distribution remains constant as *isomodular*

[1] The Greek for "to penetrate" is "diabaínein."

—a term which is not in general use. We may realize a process of this kind by supposing the system to be connected to a *reservoir* of the modulus in question throughout the duration of the change of state.

Since the quantities $E$ and $W$ depend only on the instantaneous state of the system, especially simple relations result for the case in which, at the conclusion of the changes of state, a system finally returns to the same definite state whence it started. A process of this kind, wherein the initial and final states are identical, is called a *cyclic process*. From eqns. (1) and (2) we obtain for such a process the simple, but none the less important, relations

$$\sum dQ = \sum dA\,, \tag{5}$$

and

$$\sum \frac{dQ}{\Theta} = 0\,. \tag{6}$$

We shall now consider a special cyclic process of great importance, composed of two isomodular and two adiabatic component processes. The nature of such a process can best be illustrated with the help of a *graphical representation*. Owing to the existence of the equation of state, each state of the system considered can be described by two variables. We can therefore represent each state graphically by a point in a plane coordinate system, the abscissa of the point being equal to the value of the volume, and the ordinate to the value of the pressure. Finite *changes of state* are then represented by *curves*, the successively arranged points on which indicate the consecutive states through which the system passes during the process. Curves representing isomodular or adiabatic changes of state will, for the sake of brevity, be called isomodular or adiabatic curves.

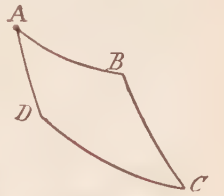


FIG. 125.

Let us now consider a definite state, graphically represented by the point $A$ (Fig. 125). We allow the system to expand isomodularly (the volume-abscissæ therefore increasing) to a second state, represented by the point $B$. To this change of state will correspond the segment $AB$ of an isomodular curve. Let the second component process to which the system is subjected be an adiabatic expansion. This will lead to a third state, represented by the point $C$ (with a

still larger abscissa), the interval $BC$ being a segment of an adiabatic curve. We draw now an isomodular curve through the point $C$, on the one hand, and on the other an adiabatic curve through the point $A$. Let the two curves intersect at a point $D$,[2] representing a fourth state. The third component process will then consist in an isomodular passage of the system from the third to the fourth state, and the fourth component process in an adiabatic return from the fourth to the original first state.

Let $Q_1$ and $Q_2$ be the quantities of energy taken up during the first and third isomodular processes (it is hardly necessary to point out that $Q_2$ will obviously be negative). The quantity of energy $Q_1$ will be taken up at a modulus value of $\Theta_1$, and the quantity $Q_2$ at a modulus value of $\Theta_2$. Since the second and fourth component processes are adiabatic, the total quantity of energy introduced into the system from without during the cycle is equal to $(Q_1 + Q_2)$. Applying eqns. (5) and (6) to this special case of a cyclic process, we obtain therefore the two relations

$$Q_1 + Q_2 = \sum dA\,, \tag{7}$$

and

$$\frac{Q_1}{\Theta_1} + \frac{Q_2}{\Theta_2} = 0\,. \tag{8}$$

If we denote the sum of $dA$ briefly by $A$, meaning thereby the total work performed by the system during the cycle,[3] we have, by eqns. (7) and (8),

$$A = Q_1 + Q_2 = Q_1 - Q_1 \frac{\Theta_2}{\Theta_1} = Q_1 \frac{\Theta_1 - \Theta_2}{\Theta_1}\,. \tag{9}$$

Thus, of the quantity of energy taken up by the system during the first component process, *only a fraction* $(\eta\ Q_1)$ is *transformed into work*. The quantity $\eta$ is called the *efficiency of the cycle*, and is given by the relation

$$\eta = \frac{\Theta_1 - \Theta_2}{\Theta_1}. \tag{10}$$

The efficiency of the cycle considered is therefore *quite independent of all properties of the system* subjected to the cyclic

[2] It is admittedly taken for granted that, starting from the state $A$, the modulus *can* be brought, by an adiabatic change of state, to the value corresponding to the point $C$.

[3] Work performed by the system is counted as positive, and work expended on the system as negative. For the rest, $dA$ is obviously not the differential of $A$ ; *cf.* § 148, footnote 1.

process. It depends simply and solely on the *values of the moduli* associated with the two reservoirs in the cycle.

This last fact, now, enables us to construct a *scale of modulus values* without our requiring to know anything about the nature of the internal energy of the systems. It is sufficient for us to be able to determine the quantities of energy transferred to or from the systems with respect to external sources, their volume and their pressure, and the work performed when the volume is changed. Apart from this, we require to have only two systems, which we may call $S_1$ and $S_2$, in quite definite states (easily identifiable and reproducible). These states, which we shall regard as *fundamental*, must evidently be associated with two definite values of the modulus of distribution ($\Theta_1$ and $\Theta_2$).

We may, quite *arbitrarily*, take the modulus of distribution of the system $S_2$ as the *zero-point of a scale*, and, likewise arbitrarily, *define a scale-unit* by making the difference between the moduli of distribution of the systems $S_1$ and $S_2$ equal to $a$ scale-units. Then we carry out a cyclic process between the systems $S_1$ and $S_2$ with the aid of an auxiliary system with respect to which those systems act as reservoirs; let the measured efficiency of this process be $\eta^*$. This gives us at once, however, the number of scale-units by which the *absolute zero* of modulus values lies beneath the arbitrary zero. For, if this number be $\zeta$, we have

$$(\Theta_1 - \Theta_2) : \Theta_1 = a : (a + \zeta),$$

and hence, by eqn. (10),

$$\zeta = a \left( \frac{1}{\eta^*} - 1 \right). \tag{11}$$

If, now, any system ($S'$) be given, we can find the exact value ($z$) on the scale corresponding to its modulus value ($\Theta'$) by carrying out a cycle between the systems $S'$ and $S_2$. Let the measured efficiency be $\eta'$; then

$$\eta' = \frac{\Theta' - \Theta_2}{\Theta'},$$

or,

$$\eta' = \frac{z}{z + \zeta}; \tag{12}$$

and, since $\zeta$ is known already, $z$ can be calculated.

Thus, it is possible for an observer to prepare an exact scale of modulus values even when the interior of the system is completely hidden from him, and he can only follow the behaviour of the system from outside.

## § 152. Theorem of Equipartition.

Apart from the quantities that describe the state of a system as a whole (modulus, parameters, statistical forces), we can also consider those quantities which determine the state of the individual. Let us investigate the *special case* in which the *energy of the particles* is *purely kinetic*, so that, as is well known, it may be represented in the form

$$\epsilon_i = \frac{1}{2m}(p_x^2 + p_y^2 + p_z^2), \tag{1}$$

where $p_x$, $p_y$, $p_z$ denote the components of the momentum. We shall suppose the components of the momentum to be *graduated* in definite *intervals*, and we shall put

$$\Delta p_x \, \Delta p_y \, \Delta p_z = \Delta\tau . \tag{2}$$

We find then

$$e^{-\frac{\epsilon_i}{\Theta}} = e^{-\frac{p_x^2}{2m\Theta}} e^{-\frac{p_y^2}{2m\Theta}} e^{-\frac{p_z^2}{2m\Theta}} . \tag{3}$$

Now, on multiplying the right-hand side of eqn. (3) by $dp_x, dp_y, dp_z$ and *integrating* over all values of $p_x, p_y, p_z$, we shall evidently obtain the sum of $e^{-\epsilon_i/\Theta}$ over all individual states, multiplied by $\Delta\tau$. Hence we have

$$\sum e^{-\frac{\epsilon_i}{\Theta}} = \frac{1}{\Delta\tau} P^3 , \tag{4}$$

where, for brevity,

$$P = \int_{-\infty}^{+\infty} e^{-\frac{p_x^2}{2m\Theta}} dp_x = \sqrt{2m\Theta} \int_{-\infty}^{+\infty} e^{-\frac{p_x^2}{2m\Theta}} d\left(-\frac{p_x}{\sqrt{2m\Theta}}\right). \tag{5}$$

We know, however, that for a quite arbitrary variable $\xi$, the definite integral[1]

[1] Eqn. (6) is most simply derived as follows :—Let the (as yet unknown) value of the definite integral be $J$. If we suppose an arbitrary plane coordinate system to be constructed with a $\xi$- and an $\eta$-axis, then we must have

$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-(\xi^2+\eta^2)} d\xi \, d\eta = J^2 .$$

On introducing polar instead of rectangular coordinates, we have

$$\xi^2 + \eta^2 = r^2 ,$$

and the surface-element $d\xi . d\eta$, with the integration limits $-\infty$ and $+\infty$, is replaced by the surface-element $r\,d\phi\,.\,dr$ with the integration limits 0 and $\infty$ for $r$, and 0 and $2\pi$ for $\phi$. Hence

$$J^2 = \int_0^{\infty}\int_0^{2\pi} e^{-r^2} r\,dr\,d\phi = 2\pi\int_0^{\infty} e^{-r^2} r\,dr = \pi\int_0^{\infty} e^{-r^2} d(r^2) = \pi ,$$

and therefore

$$J = \sqrt{\pi} .$$

$$\int_{-\infty}^{+\infty} e^{-\xi^2} d\xi = \sqrt{\pi} . \tag{6}$$

Therefore,

$$\sum e^{-\frac{\epsilon_i}{\Theta}} = \frac{1}{\Delta\tau} (2\pi m \Theta)^{\frac{3}{2}} . \tag{7}$$

If we differentiate this equation partially with respect to $\Theta$, we find

$$\frac{1}{\Theta^2} \sum \epsilon_i e^{-\frac{\epsilon_i}{\Theta}} = \frac{1}{\Delta\tau} \cdot \frac{3}{2} \cdot \Theta^{\frac{1}{2}} (2\pi m)^{\frac{3}{2}} . \tag{8}$$

On multiplying this equation by $\Theta^2$ and then dividing by eqn. (7), we obtain on the left-hand side [in accordance with § 147 eqn. (16)] the *mean value of the individual energy*, which may be denoted by $\bar{\epsilon}$. This division gives the simple relation

$$\bar{\epsilon} = \frac{3}{2} \Theta . \tag{9}$$

In the special case considered, therefore, the mean value of the individual energy is equal to *half the modulus of distribution* multiplied by the *number of degrees of freedom*. Each degree of freedom is thus associated with the same amount of energy, and, consequently, the principle expressed by eqn. (9) is known as the *theorem of equipartition*.

Our next problem is to calculate the free energy and the statistical probability for the system under consideration. The *free energy* follows at once from § 147 eqn. (18) :—

$$\Psi = - N \Theta \log_e \frac{(2\pi m \Theta)^{\frac{3}{2}}}{\Delta\tau} . \tag{10}$$

Again, according to eqn. (9), the *energy* is given by

$$E = \frac{3}{2} N \Theta . \tag{11}$$

Hence we find for the *statistical probability*, from § 147 eqn. (19),

$$\log_e W = \frac{3}{2} N + N \log_e \frac{(2\pi m \Theta)^{\frac{3}{2}}}{\Delta\tau} , \tag{12}$$

or, bearing in mind that $x$ is equal to $\log_e e^x$,

$$\log_e W = N \log_e \frac{(2\pi e m \Theta)^{\frac{3}{2}}}{\Delta\tau} . \tag{13}$$

Finally, we shall investigate the number of particles, the

*velocity components* of which lie within definite *intervals*, *e.g.*, between $v_x$ and $(v_x + \Delta v_x)$, *etc.*, where, in the sense of eqn. (2)

$$m\,\Delta v_x \,.\, m\,\Delta v_y \,.\, m\,\Delta v_z = \Delta\tau\,. \tag{14}$$

Let $n'$ be the number of these particles. Then, by § 147, eqn. (12),

$$n' = N\,w_i = N\,\frac{e^{-\frac{\epsilon_i}{\Theta}}}{e^{-\frac{\psi}{\Theta}}}\,. \tag{15}$$

But, in accordance with eqn. (10),

$$e^{-\frac{\psi}{N\Theta}} = \frac{(2\,\pi\,m\,\Theta)^{\frac{3}{2}}}{\Delta\tau}\,. \tag{16}$$

Therefore,

$$n' = N\left(\frac{m}{2\,\pi\,\Theta}\right)^{\frac{3}{2}} e^{-\frac{m}{2\,\Theta}(v_x^2 + v_y^2 + v_z^2)}\,\Delta v_x\,\Delta v_y\,\Delta v_z\,. \tag{17}$$

This equation, in which $\frac{3}{2}\,\Theta$ represents the *mean value* of the energy of a single particle, is an expression of the famous *Theorem of Distribution of Velocities*, which was first derived in 1860 by *Maxwell*.

## § 153. The Theorem of the Virial.

A statistical relation of importance for systems of *many particles* is obtained when we multiply the equations of motion of the free particle,

$$X = m\,\frac{d^2x}{dt^2}\,,\ etc., \tag{1}$$

by the corresponding coordinates, and add them together. We find in the first place

$$Xx + Yy + Zz = m\left[x\,\frac{d^2x}{dt^2} + y\,\frac{d^2y}{dt^2} + z\,\frac{d^2z}{dt^2}\right]. \tag{2}$$

On the other hand, however,

$$x\,\frac{d^2x}{dt^2} = x\,\frac{dv_x}{dt} = \frac{d}{dt}\,(x\,v_x) - {v_x}^2 = \frac{d^2}{dt^2}\left(\frac{x^2}{2}\right) - {v_x}^2\,. \tag{3}$$

Now, following *Clausius*, we define as the *virial* of a mass system the expression

$$U = -\,\frac{1}{2}\sum_{h=1}^{h=n}(X_h\,x_h + Y_h\,y_h + Z_h\,z_h)\,. \tag{4}$$

If, therefore, we form eqn. (2) for all the $n$ particles in the system and write down the sum, we obtain

$$(5) \qquad -U = \frac{1}{4}\sum m \frac{d^2}{dt^2}(x^2 + y^2 + z^2) - L\,,$$

where $L$ denotes the kinetic energy.

Finally, let us multiply eqn. (5) by $dt$, integrate over a sufficiently large time $T$, and divide by this time. The left-hand side of the equation will then give the *mean value* of the virial *with respect to the time* (generally known simply as the virial), and the last term in the equation will represent the mean value of the kinetic energy over the same period. The first term on the right-hand side, however, will yield the expression

$$(6) \qquad \frac{1}{T}\sum m \frac{d}{dt}(x^2 + y^2 + z^2)\,,$$

and this expression becomes *arbitrarily small* if a sufficiently large value be assigned to $T$. Thus eqn. (5) represents the *theorem of the virial*, as developed by *Clausius* in 1870, according to which the mean value of the kinetic energy of a system is equal to its virial.

To take an extremely simple *example*, we shall calculate the virial of a system consisting of uniform particles occupying a space enclosed by a parallelepipedal container. We assume the absence of any interactional forces between the particles, the only forces exerted being those between the containing walls and the particles which collide with them. We shall denote the lengths of the sides of the container by $a$, $b$, $c$, and imagine a coordinate system to be constructed with its axes parallel to the three sides. The mean value, with respect to the time and to the unit of surface, of the force exerted by the walls on the system is equal and opposite to the *pressure* ($p$) of the system. Hence

$$(7) \qquad \sum X_h x_h = -p\,b\,c\,(x_2 - x_1)\,,$$

where $x_1$ and $x_2$ are the distances separating the two surfaces parallel to the $y$-$z$-plane from the latter. Since $(x_2 - x_1)$ is equal to $a$, however, we have, denoting the volume of the parallelelepiped by $V$,

$$(8) \qquad \sum X_h x_h = -p\,V\,.$$

The same value is obtained for $\sum Y_h y_h$ and for $\sum Z_h z_h$, and therefore it follows from eqn. (4) that, in the case under consideration,

$$U = \frac{3}{2} p V . \tag{9}$$

On the other hand, according to the theorem of equipartition, inasmuch as there is here no potential energy, the kinetic energy of the system is given by

$$L = \frac{3}{2} N \Theta , \tag{10}$$

where $N$ is the number of particles, and $\Theta$ the modulus of distribution. Equating the right-hand sides of eqns. (9) and (10) in the sense of the virial theorem, we obtain, therefore, the important relation

$$p V = N \Theta . \tag{11}$$

### § 154. Rayleigh and Jeans's Principle.

In addition to the *special case of a distribution of energy* considered in the previous sections, there is another of the

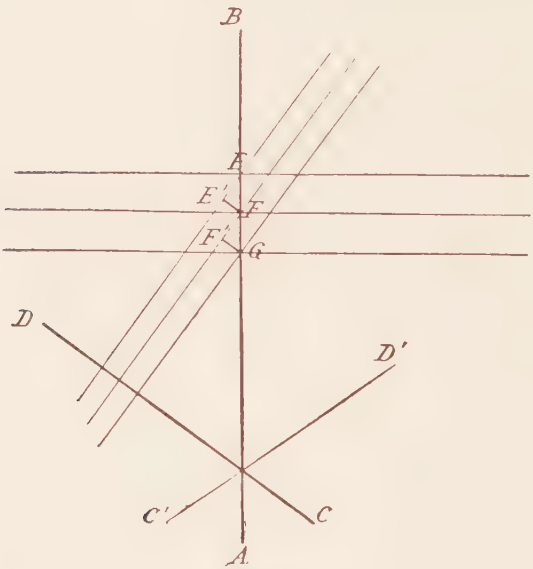



Fig. 126.

greatest importance for theoretical physics. This concerns the *canonical distribution of quanta of oscillatory energy*. If we consider an enclosed space, which, for simplicity's

sake, we shall assume to be a *cube*, it may contain *stationary vibrations*, so-called *characteristic vibrations*, of the most varied kinds. Any investigation of the energy distribution must first be preceded by a solution of the question as to how these vibrations are distributed over the various frequency values. If we graduate the frequency values in intervals, we must first find out how large is the *number of vibrations*, the *frequencies* of which *lie within a definite interval*, for a given volume of cube.

In answering this question, we begin with the well-known fact that, along a stretched string of length $l$, only such stationary vibrations are possible as have wave-lengths ($\lambda$), which, for integral values of the *number of loops* ($n$), satisfy the equation

$$\frac{\lambda}{2} = \frac{l}{n}. \tag{1}$$

In a cube the nodal points are replaced by *nodal planes* parallel to the faces of the cube, and the values of the segmental numbers may differ for the three mutually perpendicular directions parallel to the edges. We shall imagine a coordinate system to be constructed with a corner of the cube as origin, and the three edges meeting at this corner as coordinate axes. Let the number of segments in the $x$-, the $y$-, and the $z$-direction be denoted by $n_1$, $n_2$, $n_3$, and the angles included between an arbitrary ray and the three axes by $\alpha$, $\beta$, $\gamma$. If, now, the straight line $AB$ be parallel to the $z$-axis (Fig. 126), and if $CD$ be the direction of a ray, then, when the nodal planes parallel to the $x$-$y$-plane intersect $AB$ at $E$, $F$, $G$ at intervals of $a/n_3$, the wave-planes passing through $E$, $F$, $G$ must also exhibit differences in phase of half a period. We must therefore have

$$E'F = F'G = \lambda/2.$$

But

$$E'F = EF\cos\gamma = \frac{a}{n_3}\cos\gamma,$$

and hence

$$\frac{\lambda}{2} = \frac{a}{n_3}\cos\gamma.$$

Since analogous relations must hold for the $x$- and the $y$-axis, we obtain a triplet of equations :

$$(2)\qquad \begin{cases} \frac{2a}{\lambda}\cos\alpha = n_1, \\ \frac{2a}{\lambda}\cos\beta = n_2, \\ \frac{2a}{\lambda}\cos\gamma = n_3. \end{cases}$$

Again, according to a well-known formula,

$$(3)\qquad \cos^2\alpha + \cos^2\beta + \cos^2\gamma = 1.$$

Hence, on squaring the three eqns. (2) and adding, we find

$$(4)\qquad \frac{2a}{\lambda} = \pm\sqrt{{n_1}^2 + {n_2}^2 + {n_3}^2}.$$

Inserting this value in eqns. (2), we obtain the relations

$$(5)\qquad \begin{cases} \cos\alpha = \pm\frac{n_1}{\sqrt{{n_1}^2 + {n_2}^2 + {n_3}^2}}, \\ \cos\beta = \pm\frac{n_2}{\sqrt{{n_1}^2 + {n_2}^2 + {n_3}^2}}, \\ \cos\gamma = \pm\frac{n_3}{\sqrt{{n_1}^2 + {n_2}^2 + {n_3}^2}}. \end{cases}$$

(In Fig. 126, therefore, in addition to the direction $CD$, the symmetrical direction $C'D'$ would be possible.)

Thus we see that to *every numerical triplet*, $n_1, n_2, n_3$, there corresponds a *characteristic vibration of quite definite wave-length and quite definite direction*. Conversely, *only* those vibrations are possible for which the values of $\lambda$, $\alpha$, $\beta$, $\gamma$ satisfy eqns. (4) and (5) for integral values of $n_1, n_2, n_3$.

If we replace the wave-length in eqn. (4) by the frequency ($\nu$) and the wave-velocity ($v$), according to the well-known formula

$$\frac{1}{\lambda} = \frac{\nu}{v}$$

[§ 42 eqn. (9)], we may also write eqn. (4) in the form

$$(6)\qquad \frac{2a}{v}\nu = \sqrt{{n_1}^2 + {n_2}^2 + {n_3}^2}$$

Let us now consider a quite arbitrary *vibration interval*, extending from a frequency $\nu'$ to a frequency $\nu''$ (differing but slightly from $\nu'$). We shall denote the difference $(\nu'' - \nu')$ by $\Delta\nu$. The question then arises as to how large, in view of eqn. (6), is the possible number of frequencies which lie within the selected interval.

To answer this question, we may conveniently make use of a simple *geometrical construction*. We imagine a rectangular coordinate system to be constructed, using, however, only the positive halves of the axes, so that an *octant* is cut out of the whole space. We suppose then a *spatial point-lattice* to be constructed in this octant, with the unit of length as lattice constant. The individual points in the lattice will thus everywhere possess integral coordinates; and, just as each point is associated with a numerical triplet $(n_1, n_2, n_3)$ as coordinate values, so, conversely, each numerical triplet $(n_1, n_2, n_3)$ unequivocally determines a point in the lattice. Consequently, each lattice point determines a stationary vibration of given frequency and direction, while, conversely, every such vibration has a representative in a quite definite point in the lattice.

Now the distance between a lattice point and the origin of the octant is given by

$$r = \sqrt{n_1^2 + n_2^2 + n_3^2}, \tag{7}$$

or, by eqn. (6),

$$r = \frac{2a}{v}\nu, \tag{8}$$

where $\nu$ is taken to mean the frequency of the characteristic vibration corresponding to the lattice point concerned.

The *number of characteristic vibrations*, with frequencies between $\nu'$ and $\nu''$, which are possible within the cube is therefore *equal simply to the number of lattice points* lying within the space bounded by two spherical surfaces with radii $r'$ and $r''$, where

$$r' = \frac{2a}{v}\nu', \qquad r'' = \frac{2a}{v}\nu''. \tag{9}$$

Since, on an average, each unit of volume contains one lattice point (because the distance between two neighbouring points is equal to the unit of length), the number of points within the spatial octant bounded by the two spherical surfaces is approximately equal to one-eighth of the volume of the whole spherical shell enclosed by the two concentric surfaces. We shall assume—as will always be the case in subsequent considerations—that $\nu'$ and $\nu''$ are extremely large numbers [1]; then, even when $\Delta\nu$ is small compared with $\nu'$, the corresponding difference $(r'' - r')$ can be regarded as large in comparison with the unit of length. The

[1] In optics the frequency is of the order of magnitude of $10^{14}$.

volume of an eighth of the spherical shell will thus be approximately given by

$$V' = \frac{1}{8} 4\pi r'^2 (r'' - r'). \tag{10}$$

The number of possible characteristic vibrations in the interval between $\nu'$ and $\nu''$ is equal in magnitude to this volume. Denoting this number by $\Delta Z$, we have [in virtue of eqn. (9) ]

$$\Delta Z = \frac{\pi}{2}\left(\frac{2a}{v}\nu'^2\right).\frac{2a}{v}.\Delta\nu,$$

or, replacing $\nu'$ simply by $\nu$,

$$\Delta Z = \frac{4\pi a^3 \nu^2}{v^3}\Delta\nu. \tag{11}$$

Since $a^3$ represents the volume of the cube, the quantity

$$z_\nu = \frac{4\pi\nu^2}{v^3}, \tag{12}$$

referred to the unit of volume, may be defined as the *density of the vibrations in the frequency interval between* $\nu$ *and* $(\nu + d\nu)$. On multiplying $z_\nu$ by the magnitude of the interval and the volume, we obtain the number of vibrations, the frequencies of which lie within the interval. Eqn. (12), which shows that the *vibration density is proportional to the square of the frequency*, was first derived by *Rayleigh* (1900), and more accurately established by *Jeans* in 1905.[2]

## § 155. The Statistical Equilibrium of a Vibrating System.

Let us again consider a cube filled with stationary vibrations, and fix our attention upon those of which the frequencies lie in the very small interval between $\nu$ and $(\nu + d\nu)$. These vibrations of *one and the same frequency* $\nu$ may be regarded as an assembly of individuals, the energy of which is assumed to be canonically distributed. According to the *quantum principle*, every characteristic vibration can then occur in a series of *discrete states* in which the associated values of the energy will be $0, h\nu, 2h\nu, 3h\nu$, *etc.* If $\Theta$ be the modulus of distribution, we have thus

$$\sum e^{-\frac{\epsilon_i}{\Theta}} = \sum_{n=0}^{n=\infty} e^{-\frac{nh\nu}{\Theta}}. \tag{1}$$

[2] Admittedly, in this section, only the special case of a cube has been considered. But it will be evident from the considerations of the next section, and even more so from subsequent applications, that eqn. (12) must hold quite independently of all geometrical shape.

In accordance with the well-known summation formula for a geometrical progression,

$$1 + y + y^2 + y^3 + \ldots = \frac{1}{1-y}, \tag{2}$$

when $y < 1$. Hence

$$\sum e^{-\frac{\epsilon_i}{\Theta}} = \frac{1}{1 - e^{-\frac{h\nu}{\Theta}}}. \tag{3}$$

On differentiating partially with respect to $\Theta$, and then multiplying both sides of the equation by $\Theta^2$, we find

$$\sum \epsilon_i e^{-\frac{\epsilon_i}{\Theta}} = \frac{h\nu \cdot e^{-\frac{h\nu}{\Theta}}}{\left(1 - e^{-\frac{h\nu}{\Theta}}\right)^2}. \tag{4}$$

If we now divide eqn. (4) by eqn. (3), it follows from § 147 eqn. (16) that the left-hand side must represent the *mean value of the energy* associated with a single vibration of frequency $\nu$. Thus

$$\bar{\epsilon} = \frac{h\nu}{e^{\frac{h\nu}{\Theta}} - 1}. \tag{5}$$

On multiplying this value by the density of the vibrations of frequency $\nu$ [§ 154 eqn. (12)], we obtain the *energy density* for the frequency considered, and this may be denoted by $u_\nu$.[1] Hence, we have

$$u_\nu = \frac{4\pi h\nu^3}{v^3} \cdot \frac{1}{e^{\frac{h\nu}{\Theta}} - 1}, \tag{6}$$

where $v$ again denotes the velocity of propagation of the vibrations.

The whole system of stationary vibrations can be resolved into *component systems* of definite frequency values if we graduate these values into intervals. We have hitherto considered such a component system, composed of vibrations of the same frequency. If, however, the whole system of vibrations is to be *in statistical equilibrium*, it follows from § 149 that this is only possible when *all the component systems, despite the different frequencies*, have the *same value of modulus of distribution*.

[1] It must be remembered, however, that the physical dimensions of $u_\nu$ are not those of an energy density. That is true only for the product of $u_\nu$ and $d\nu$.

When such is the case, the true *energy density*, denoted by $\eta$, may be calculated from the formula

$$\eta = \int_0^\infty u_\nu \, d\nu , \tag{7}$$

$\Theta$ being regarded as constant during the integration. Hence

$$\eta = \frac{4\pi h}{v^3} \int_0^\infty \frac{\nu^3 \, d\nu}{e^{\frac{h\nu}{\Theta}} - 1} . \tag{8}$$

To simplify this expression, let us put

$$\frac{h\nu}{\Theta} = x , \qquad \text{or } \nu = \frac{x\Theta}{h} : \tag{9}$$

then

$$\eta = \frac{4\pi \Theta^4}{h^3 v^3} \int_0^\infty \frac{x^3 \, dx}{e^x - 1} . \tag{10}$$

We can evaluate the definite integral in this equation by starting off with eqn. (2), in which we replace $y$ by $e^{-x}$: then

$$\frac{1}{1 - e^{-x}} = \frac{e^x}{e^x - 1} = 1 + e^{-x} + e^{-2x} + e^{-3x} + \ldots ,$$

or,

$$\frac{1}{e^x - 1} = e^{-x} + e^{-2x} + e^{-3x} + \ldots .$$

Hence

$$\int \frac{x^3 \, dx}{e^x - 1} = \int e^{-x} \, x^3 \, dx + \frac{1}{2^4} \int e^{-2x} \, (2x)^3 \, d(2x) + \frac{1}{3^4} \int e^{-3x} \, (3x)^3 \, d(3x) + \ldots .$$

Now, according to a well-known formula [2] for an arbitrary variable $z$,

$$\int_0^\infty z^3 e^{-z} \, dz = 3 ! . \tag{11}$$

[2] This formula may be derived as follows :—For an arbitrary parameter $t$

$$\int_0^\infty e^{-tz} \, dz = \left| \frac{1}{t} \cdot e^{-tz} \right|_\infty^0 = \frac{1}{t} .$$

On successive partial differentiations with respect to $t$, we find

$$\int_0^\infty z \, e^{-tz} \, dz = \frac{1}{t^2} ; \quad \int_0^\infty z^2 e^{-tz} \, dz = \frac{1 \cdot 2}{t^3} ; \quad \int_0^\infty z^3 e^{-tz} \, dz = \frac{1 \cdot 2 \cdot 3}{t^4} . \tag{11b}$$

When $t$ is put equal to 1, we obtain eqn. (11).

Since 3 ! is equal to 6, we have

$$\int_0^\infty \frac{x^3\,dx}{e^x - 1} = 6\,\alpha, \tag{12}$$

where $\alpha$ denotes the sum of the convergent power series [3]

$$\frac{1}{1^4} + \frac{1}{2^4} + \frac{1}{3^4} + \frac{1}{4^4} + \ldots = 1{\cdot}0823. \tag{13}$$

Hence

$$\eta = \frac{24\,\pi\,\alpha\,\Theta^4}{h^3\,v^3}. \tag{14}$$

In statistical equilibrium, therefore, the *energy density is proportional to the fourth power of the modulus of distribution.*

In order next to calculate, in like manner, the *free energy* of the quantized vibrating system, we shall first consider the component system of those vibrations with frequencies lying in the interval between $\nu$ and $(\nu + d\nu)$. It follows from § 147 eqn. (18) that the amount of free energy associated with a single individual is given by

$$\psi = -\,\Theta \log_e \sum_{n=0}^{n=\infty} e^{-\frac{nh\nu}{\Theta}},$$

or, by eqn. (3),

$$\psi = \Theta \log_e \left(1 - e^{-\frac{h\nu}{\Theta}}\right). \tag{15}$$

We introduce now the well-known logarithmic series $(y < 1)$

$$\log_e (1 - y) = -\,y - \frac{y^2}{2} - \frac{y^3}{3} - \frac{y^4}{4} - \ldots. \tag{16}$$

Then, making use of the abbreviation given in eqn. (9), we find

$$\psi = -\,\Theta\left[e^{-x} + \frac{e^{-2x}}{2} + \frac{e^{-3x}}{3} + \frac{e^{-4x}}{4} + \ldots\right]. \tag{17}$$

On multiplying this quantity by the density of the vibrations in the frequency interval between $\nu$ and $(\nu + d\nu)$, and then integrating with respect to $\nu$ from 0 to $\infty$, we obtain the *density of the free energy*, which may be denoted by $\zeta$. Thus,

$$\zeta = -\,\frac{4\,\pi\,\Theta}{v^3}\int_0^\infty \left[e^{-x} + \frac{e^{-2x}}{2} + \ldots\right]\nu^2\,d\nu,$$

or, by eqn. (9),

[3] Although it does not concern the subsequent considerations at all, we may remark in passing that $\alpha$ is actually $\pi^4 / 90$.

$$(18)\quad \left\{\begin{aligned}\zeta = -\frac{4\pi\Theta^4}{h^3 v^3}\Big[&\int_0^\infty e^{-x} x^2\, dx + \frac{1}{2^4}\int_0^\infty e^{-2x}(2x)^2\, d(2x)\\ &+\frac{1}{3^4}\int_0^\infty e^{-3x}(3x)^2\, d(3x)+\dots.\Big].\end{aligned}\right.$$

But, for any arbitrary variable ($z$),

$$(19)\qquad \int_0^\infty e^{-z} z^2\, dz = 2$$

[according to eqn. (11b) ]. Hence, introducing the abbreviation used in eqn. (13),

$$(20)\qquad \zeta = -\frac{8\pi\alpha\Theta^4}{h^3 v^3}.$$

As we see by a comparison with eqn. (14), the *free energy* is therefore *equal*, but opposite in sign, *to one-third of the total energy*. We have

$$(21)\qquad \Psi = -\frac{E}{3}.$$

## Chapter XIX

# THERMODYNAMICS

### § 156. Temperature Scales

A THEORETICAL explanation is afforded by statistics for those phenomena which are known as heat phenomena, if, in every case in which *energy* is *distributed among a very large number of individuals* in a manner not directly perceptible to human senses, this energy be *interpreted as heat.* The individuals among which heat energy is distributed *throughout bodies* are identified by theoretical physics with *molecules.*[1]

All quantities which, in a statistical sense, are proportional to the number of individuals must therefore, in a thermal sense, be *proportional to the number of molecules*, or, what amounts to the same thing in chemically homogeneous substances, to the *mass.* Thus all these quantities may be *referred* to a *gram-molecule*, or to the *unit of mass.*

It follows from the electron theory that the forces acting between the particles of matter must be regarded as central forces and depend only on the distances between the particles. According to § 17, however, a system of this kind is governed by the Principle of the Conservation of Energy, which was one of our hypotheses in our previous statistical considerations of energy. If, now, the internal forces only depend on the distances separating the individual molecules, then the internal energy of a homogeneous body will evidently depend only on one parameter, *viz.*, on the volume of the unit of mass, or the so-called *specific volume.* It seems justifiable, therefore, to regard bodies of a *physically uniform nature* as being also *statistically homogeneous.*

Experimental physics is acquainted with various phenomena in which *energy* arising from the *motion of bodies*, or *electromagnetic energy, apparently disappears with evolution of heat.* This is the case when energy of motion is destroyed by any obstacles such as *friction* or the resistance of the

[1] In elements which do not form molecules, the atoms are correspondingly regarded as "monatomic molecules."

medium ; when an electric current is transmitted through a conductor possessing a *galvanic resistance ;* when electromagnetic waves strike a body which does *not completely reflect them.* In all these instances, theoretical physics assumes that the energy which has apparently disappeared is not really lost, but is distributed among a large number of individuals. The *heat* produced in such cases is therefore *measurable in units of energy,* if the apparent loss of energy be determined, as may easily be done with frictional heat and electrical heat.

According to this concept, the energy introduced into or abstracted from a system of molecules appears as *heat absorbed* or *emitted.* Now if, on connecting two bodies in such a way that a heat transfer becomes possible, no change in their initial state occurs, it must be concluded from the existence of the equation of state that both bodies have the *same modulus of heat.* A body of which the state is readily variable may thus, to a certain extent, serve as a thermoscope, *viz.*, as an *instrument for the detection of the equality of heat moduli.* For, if the volume at constant pressure, or the pressure at constant volume, indicated by this instrument be the same when it is connected with two bodies $A$ and $B$, we may conclude that, at the moment when they were connected with the instrument, the bodies $A$ and $B$ had the same heat modulus. Furthermore, if the change of state of a body $K$ occurs in such a way that the state of a thermoscope connected with the body does not vary throughout the change, we know that the change of state of $K$ is an *isomodular* one. On the other hand, we can effect an *adiabatic* change of state in a body by surrounding it completely with some substance which is known from experience to be impermeable to heat (as is very approximately the case with, for instance, wool).

Measurements made on *cyclic processes* based on these facts permit of the establishment of an *exact scale of values of heat moduli.* Let us suppose a homogeneous body (preferably a gas, because of its ready variability of state) to be in a state I, and a thermoscope connected with it to give a reading $a$. Before commencing the cyclic process, let the volume be adiabatically increased until the thermoscope gives any reading $b$, the state of the body, as defined by its volume and pressure, being IV. Then we have to carry out a cyclic process in the following manner :—

We transfer the body from state I to an arbitrary state (II), by means of an isomodular expansion, and then we let it expand adiabatically to a state III in which the thermo-

scope gives the above-mentioned reading $b$. Next, the body is brought to state IV, by means of an isomodular compression, and finally adiabatically transferred to state I. We measure in units of energy, on the one hand, the quantity of heat introduced, for example, electrically by means of a suitable regulation of current during the isomodular transference from I to II, and, on the other hand, the total work performed during the cycle. This may most simply be done by representing the process graphically on a $p$-$V$-plane, and measuring the area of the figure bounded by the four component curves (the so-called indicator diagram), for this area is equal to $\int p\,dV$, or the work performed. On dividing this work by the above quantity of heat (in energy units), we obtain the efficiency of the cycle.

We introduce, now, as *fundamental values* of the heat modulus the values possessed by *melting ice and boiling water at normal atmospheric pressure*,[2] it being empirically established (from reasons which will appear later [3]) that the reading of the thermoscope does not change during either the melting or the boiling process. In conformity with experimental physics, we choose the ice-point as the zero-point on an arbitrary scale, and we divide the difference between the two fundamental values of the modulus arbitrarily into *one hundred* scale-units or *degrees*. Then we define the number on this arbitrary scale corresponding to the heat modulus as the *temperature*.

If we carry out a cycle in such a way that the isomodular component processes are performed at 100° and 0°, we can determine [according to § 151 eqn. (11)] how many degrees below the ice-point the *absolute zero of temperature* lies. As is found from other observations which are easier to make and will be discussed later,[4] this number amounts to 273·1°. The number of degrees obtained by adding the number 273·1 to that reckoned from the ice-point is known as the *absolute temperature*; it is generally denoted by the letter $T$. Modulus of heat and absolute temperature can thus only differ by a universal proportionality factor.

Finally, by carrying out the cycle in such a way that, in the first component process, the thermoscope gives any

[2] Normal atmospheric pressure is equal to the force exerted, in virtue of its weight, by a column of mercury (in thermal equilibrium with melting ice) 76 cm. in height and 1 cm.$^2$ in cross-section, at sea-level in the geographical latitude of 45°. In absolute measure, this pressure is equal to 1,013,250 gm. cm.$^{-1}$ sec.$^{-2}$.

[3] *Vide* § 163.

[4] *Vide* § 165.

arbitrary reading, while the third component process is performed at the temperature of melting ice, we can derive the correct value of the temperature for any given reading of the thermoscope, with the help of § 151 eqn. (12). Used in this manner, the thermoscope becomes an exact measurer of temperature—a *thermometer*.

---

The development of the construction of temperature scales in practical physics has followed along much simpler lines than the method given above, which must be regarded as the most perfect from the point of view of theoretical physics. If $V_0$ be the volume of a thermometric substance (such as mercury) at the ice-point, $V_{100}$ the volume at the boiling-point, and $V$ the volume at an unknown temperature $\theta$, $\theta$ is simply equated to $(V - V_0)/(V_{100} - V_0)$. The quantity $\theta$ thus appears as the temperature in the scale of the thermometric substance used. The practical utility of this simple process is supported by the fact that the scales of various suitable thermometric substances (mercury, alcohol, air, *etc.*) differ but little from each other. For reasons which will appear subsequently,[5] *hydrogen* and *helium* are regarded as the most perfect thermometric substances, and the scales of other thermometric substances are therefore corrected to the hydrogen or the helium scale.[6] The idea of basing a truly accurate temperature scale upon measurements on cyclic processes is due to *William Thomson* (*Lord Kelvin*) (1855). (In § 162 we shall discuss a suitable method for the establishment of the Kelvin temperature scale.)

## § 157. Calorimetry.

The construction of a temperature scale permits of a *simple measurement of quantities of heat*, without it ever being necessary to embark upon the complicated process of a direct measurement in energy units. In accordance with the Law of the Conservation of Energy, the quantity of heat required for a finite change of state is given by the expression

$$(1) \qquad \sum dQ = E_2 - E_1 + \int_{V_1}^{V_2} p\, dV,$$

where the subscripts 1 and 2 refer to the initial and the final

[5] *Vide* § 167.

[6] It is well known from experimental physics that the scale on which the ice-point and the boiling-point are chosen as fundamental points, the fundamental interval being made equal to one hundred degrees, is due to *Celsius* (1742).

state. When, however, the change of state occurs at *constant pressure*, the integral is simply equal to $p\,(V_2 - V_1)$, and, in that case, the total amount of heat introduced depends solely on the initial and final states. It follows, therefore, from the equation of state, that the quantity of heat required to raise a given body from an initial temperature $T_1$ to a final temperature $T_2$, at constant normal atmospheric pressure, is always the *same*, and, because of the molecular structure of the heated body, this quantity of heat must also be proportional to the mass of the body.

Thus, a *unit of heat* may be arbitrarily created by defining as a *calorie* the quantity of heat required to raise 1 gram of water at constant normal atmospheric pressure from 15° to 16° (measured from the ice-point).[1] Using the same quantities of heat (which may, for instance, be generated by sending equally strong electric currents through the same resistance for the same length of time), we can at one time raise $x$ gm. of water from 15° to 16°, and at another time $y$ gm. of water from $\theta_1°$ to $\theta_2°$. We know then that $x/y$ calories are necessary to heat 1 gm. of water from $\theta_1°$ to $\theta_2°$. In this way we can find out how many calories are required to raise 1 gm. of water through one degree at any particular temperature (*e.g.*, to heat 1 gm. of water from 30° to 31°, 0·9968 cal. are required). Any carbitrary quantity of heat can thus be measured in calories, as soon as we know the rise in temperature produced by it in a given quantity of water.

The number of energy units equivalent to one calorie is known as the *mechanical equivalent of heat*.[2] It can be directly determined by heating a quantity of water by frictional or electrical means, and expressing the quantity of heat generated, on the one hand, in calories, and on the other hand (as determined by the apparent loss in energy) in energy units. The value found by experiment for the mechanical equivalent of heat is [3]

$$J = 4{\cdot}19 \times 10^7 \text{ ergs} = 427 \text{ gm.-m.} \tag{2}$$

[1] The quantity of heat required to raise 1 gm. of water from 0° to 1° is called the zero-point calorie; it is approximately 1·008 times as large as the calorie above defined. Sometimes, however, the latter is also defined as being the quantity of heat necessary to raise 1 gm. of water from 14½° to 15½°; the difference, however, is very slight. The term "mean calorie" is given to the one-hundredth part of the quantity of heat used in heating 1 gm. of water from 0° to 100°; it is very nearly equal to the true calorie (of 15°). The large calorie is understood to be equal to one thousand "small" calories, the quantity of heat then being referred to 1 kgm. instead of 1 gm.

[2] Its reciprocal value may be called the calorimetric equivalent of work.

[3] Referred to the zero-point calorie, $J = 4{\cdot}22 \times 10^7$ ergs = 430 gm.-m.

### § 158. The Two Principal Laws of Thermodynamics.

Since temperature, measured in degrees, and modulus of heat only differ by a universal proportionality factor, the theorems obtained in general statistics must be equally valid if modulus of distribution be replaced by the product of absolute temperature and a proportionality factor, and heat be substituted for energy. As applied to the theory of heat, the theorems stated at the end of § 148 form the fundamental relations of that theory, and are known as the *two Principal Laws of Thermodynamics.*

The *First Principal Law* may be expressed as follows: If an infinitesimal quantity of heat be introduced into a system, while work is simultaneously performed, the *difference between heat introduced and work performed* represents the *complete differential* of a function, the *internal energy*, which depends only on the instantaneous state of the system and, when $r$ parameters exist, is therefore determined by $(r + 1)$ *independent variables of state.*

Moreover, this statement of the First Principal Law embodies a fact which follows directly from the statistical-energetic concept of heat, *viz.*, that in all cases where heat is developed with apparent loss in energy, a *constant transformation ratio* exists between the heat measured in calories and the energy loss measured in energy units.

The nucleus of the *Second Principal Law* is contained in the conclusion derived from general statistics, that the *quotient of the infinitesimal quantity of heat introduced and the temperature* must represent the *complete differential* of a *second function*, which likewise can depend only on the instantaneous state of the system, and must therefore also be determined by $(r + 1)$ independent variables of state when the number of parameters is $r$. This function is known in thermodynamics as the *entropy*, and is thus defined by the equation

$$dS = \frac{dQ}{T}. \tag{1}$$

Entropy must hence possess the same properties as the probability logarithm, to which it is proportional. *Entropy* likewise *appears to tend to a maximum*, and to remain *constant only in reversible processes.*

According to § 149 eqn. (1), the entropy of a system of bodies is equal to the *sum of the entropies* of the individual bodies, so that the entropy of a single body is proportional

to its mass.[1] If we speak of *thermodynamic* instead of statistical *equilibrium*, we see that such is only possible *between bodies at the same temperature*. When the temperatures differ, there is a tendency to *equalization*, and that in such a way that *heat is not of itself transferred from a colder to a warmer body*.

In accordance with eqn. (1), every addition of heat involves an increase in the entropy of a body. Since this applies to every process in which the energy (other than that of heat) of a closed system is transformed into heat by distribution amongst the constituent individuals, the energy contained in a closed system must apparently tend to be *gradually transformed completely into heat*. This apparently progressive transformation is known as the *dissipation*, scattering or *degradation of energy*.[2]

Denoting the universal proportionality factor between modulus of heat and absolute temperature by $k$, so that

$$\Theta = k\,T, \tag{2}$$

we have [by eqn. (1) and § 148 eqn. (16)]

$$S = k \log_e W. \tag{3}$$

The definition of entropy is thereby tacitly supplemented by the postulate that it shall not differ by any additive constant from the probability logarithm. The quantity $k$ is called *Boltzmann's constant* after Boltzmann, who founded the statistical concept of entropy in 1866; and the relation which states that the entropy is proportional to the logarithm of the statistical probability of the individual distribution of heat energy is known as *Boltzmann's entropy law*.

---

Thermodynamics first developed into a branch of theoretical physics in the middle of the 19th century. *Thermometry*,[3] however, was established as early as the beginning of the 18th century, and *calorimetry*[4] towards the end of that century. In 1822, moreover, *Fourier* created an exact theory of *heat conductivity* in his famous "Analytical Theory

[1] This follows, moreover, directly from the proportionality of the probability logarithm to the number of individuals.

[2] Energy is "degraded" because heat cannot be changed back again into other forms of which use might once again be made, for, in such a transformation back, without some kind of compensation, the entropy would be diminished.

[3] The real founder of thermometry was *Fahrenheit*.

[4] The differences in the specific heats of various bodies were first observed by *Wilke* (1772). Conspicuous in the further development of calorimetry were Black, Dulong and Petit, Lavoisier and Laplace, and Regnault.

of Heat." In 1842 *Robert Mayer* established the important principle of the existence of a *constant transformation ratio* between the amount of heat developed when an apparent loss of energy takes place and the energy which has apparently disappeared. The same principle was simultaneously established by *Joule*, though quite independently of Mayer. The establishment of this principle upon an accurate theoretical basis was chiefly due to *Helmholtz* (1847). A determination of the mechanical equivalent of heat was effected by Mayer along theoretical lines,[5] and by Joule by measurements of electrical and frictional heat.

The Second Principal Law of Thermodynamics was derived by *Clausius* (1850) and by *William Thomson* (1851) from purely phenomenological considerations. The concept of entropy is due to Clausius and that of the degradation of energy to Thomson. In the derivation of the Second Principal Law, it was possible to utilize the important results obtained as early as 1824 by *Sadi Carnot* in his researches on *thermal cyclic processes*.[6]

The middle of the 19th century also saw the establishment of the *mechanical theory of heat* which regards heat as individually distributed *energy of motion*.[7] This assumption afforded an immediate explanation of the First Principal Law, and the kinetic concept was applied with great success by *Krönig*, *Joule*, *Clausius*, and *Maxwell* to explain the properties of *gases* (*vide* § 165). To *Boltzmann* we owe, as already mentioned, the statistical interpretation of the Second Principal Law, the *relegation of irreversibility to atomistics*. His ideas have been further developed and broadened, especially by *Smoluchowski* (at the beginning of the 20th century); whilst the development of statistical methods through the treatment of *statistical mechanical problems* is principally due to *Gibbs*.

## § 159. Irreversibility.

The law of the *increase of entropy* is based on probability considerations in general statistics. It is, therefore, not a law in the strict sense of the word, but rather a *rule* which, in physical processes, is *more frequently obeyed than disobeyed*. If a certain process leads to a $z$-fold increase in the statistical

[5] *Vide* § 165.

[6] An exact mathematical form was given to Carnot's ideas by *Clapeyron* (1833).

[7] This concept was first stated in a scientific form by *Rumford* (about 1800).

probability of a system, and if we assume that this process can only take place in the one direction, being thus *irreversible*, then, for every $z$ instances in which the *assumption* is found to be *correct*, there will be only *one single* case in which it will be *contradicted by an experimental fact*. The number $z$ may thus be regarded as a *measure of the irreversibility*.

If, now, in any process the entropy increases by the amount $\Delta S$, the logarithm of the statistical probability will be increased by the amount $\Delta S/k$. The number $z$, which shows how many times the original value of the statistical probability has been raised, is then determined by the relation

$$z = e^{\frac{\Delta S}{k}}. \tag{1}$$

In order to be able to calculate $z$ for definite cases, we must, however, know the value of Boltzmann's constant ($k$); this is, as we shall find in subsequent considerations (§ 165), which we may anticipate, equal to $1{\cdot}34 \times 10^{-16}$ erg . deg.$^{-1}$.

As a very simple example of a process associated with an increase in entropy, we shall consider an *equalization of temperature*. We suppose two previously isolated bodies, differing in temperature by 1° (say, 15° and 16°), to be brought into contact with each other. Let the very small quantity of heat equivalent to 1 erg be transferred from the body which is warmer by one degree (*i.e.*, an amount not even sufficient to heat 1 mgm. of water 1/1000°).[1] The increase in the probability logarithm thereby produced is, by § 149 eqn. (2),

$$\frac{\Delta S}{k} = \frac{1\,\text{erg}}{k}\left(\frac{1}{288^\circ} - \frac{1}{289^\circ}\right) = 9 \times 10^{10}.$$

Hence, in this case, according to eqn. (1)

$$z > 10^{10^{10}}.$$

In the insignificant equalization of temperature considered, the irreversibility as measured by the number $z$ is therefore large beyond all human imagination, for, in order to write down the number $z$ in figures, we should require a strip of paper as long as the circumference of the earth. For $z$ cases in which the erg is transferred from the warmer to the colder body, there is but one instance in which the reverse process takes place.

Thus, if we make the assumption that heat is never

[1] If we suppose the two bodies to be sufficiently large, the heat quantity of one erg can be regarded as "infinitesimally small" in comparison with the heat content of either body.

spontaneously transferred from a colder to a warmer body to an observable extent (without some kind of compensation), the probability of refuting this assumption by experiment is inconceivably small.[2] Even if millions of physicists were to carry out afresh every minute throughout millions of years the experiment of bringing into contact at room temperature two bodies differing by a degree, the prospect of any one of them observing even only once a loss of heat by the colder instead of by the warmer body, to the extent of at least an erg, would be incredibly small. Inconceivably rare as is such an instance, however, its probability is still greater than zero; and, within *sufficiently long intervals of time*, which would certainly be long beyond all human conception, an observer would undoubtedly meet *at least once* with this *unimaginably rare exception*.

In the case of a system, the individual energy distribution of which is constantly changing, it may be said that *thermodynamic equilibrium* represents the *normal state;* the deviations therefrom, which are actually always present, would almost always completely escape notice. Only on occasions of the very greatest rarity would the deviations come within the limits of human perceptibility, and the further within these limits, the more seldom the occasions. Even *macroscopically* recognizable mechanical and electrical processes can occur spontaneously in a system through variations of distribution. A period of macroscopical processes of this kind, which of itself arises from a state of equilibrium and tends to return to it, would certainly represent only an *inconceivably rare and exceptional episode* in the history of the system. Moreover, however long this period might last, it would yet be unimaginably brief as compared with the time which must elapse before such a period of macroscopical processes of itself *returned* once again.

### § 160. Free Energy and Thermodynamic Potential.

In accordance with § 147 eqn. (19), *free energy* (for which, in thermodynamics, the usual symbol is $F$ instead of $\Psi$) is connected with the probability logarithm and the modulus of heat by the relation

$$E - F = \Theta \log_e W,$$

[2] The probability of the exceptional case becomes larger, however, when a smaller quantity of heat than an erg is considered, and when the temperature difference is taken as being less than a thousandth of a degree.

or [taking into account § 158 eqns. (2) and (3)] with temperature and entropy by the equation

$$F = E - TS. \tag{1}$$

Those quantities which, in statistics, are designated statistical forces are known as *thermal forces* in thermodynamics (although their physical dimensions are by no means necessarily those of a force). Denoting them again, by $U_1$, $U_2$, *etc.*, and regarding *free energy as a function of the temperature and the parameters*, we have, by § 150 eqn. (1), the relations

$$U_1 = -\frac{\partial F}{\partial u_1}, \qquad U_2 = -\frac{\partial F}{\partial u_2}, \text{ etc.} \tag{2}$$

Again, in accordance with § 148 eqn. (11),

$$\frac{\partial F}{\partial \Theta} = \frac{F}{\Theta} - \frac{E}{\Theta} = -\log_e W,$$

so that

$$-\frac{\partial F}{\partial T} = S. \tag{3}$$

It is therefore a property of the free energy, considered as a function of the temperature and the parameters, that its negative partial derivatives with respect to the parameters and the temperature give the thermal forces and the entropy respectively.

It follows from § 148 eqn. (9) that free energy and *work* performed are connected by the relation

$$dF = \frac{\partial F}{\partial T}\, dT - dA, \tag{4}$$

or [by eqn. (3) and § 150 eqn. (2)],

$$dF = -S\, dT - U_1\, du_1 - U_2\, du_2 - \ldots. \tag{5}$$

Hence, if a process be carried out at constant temperature or (as we may say by analogy with the previously used term "isomodular") *isothermally*, we have simply

$$dF = -dA. \tag{6}$$

In isothermal processes, therefore, the *work performed* is *equal to the decrease in free energy*.

Again, in processes where *no external work is performed* we have, by eqns. (4) and (3),

$$dF = \frac{\partial F}{\partial T}\, dT = -S\, dT. \tag{7}$$

Multiplying the first and third of these expressions by $T$, and bearing in mind eqn. (1), we obtain the relation

(8) $$E - F = -T\frac{dF}{dT}.$$

This equation was first established in 1882 by *Helmholtz* (to whom, moreover, the whole concept of free energy is due [1]), and is therefore generally known as Helmholtz's equation. It is of great importance on account of its applicability to all processes in which temperature changes occur without the performance of external work.[2]

Closely connected with the concept of free energy is that of *thermodynamic potential*. This is defined by the equation

(9) $$\Phi = F + U_1 u_1 + U_2 u_2 + \dots .$$

Total differentiation leads to the expression

$$d\Phi = dF + U_1 du_1 + U_2 du_2 + \dots + u_1 dU_1 + u_2 dU_2 + \dots ,$$

or, by eqn. (5),

(10) $$d\Phi = -S\,dT + u_1\,dU_1 + u_2\,dU_2 + \dots .$$

If thermodynamic potential be regarded as a *function of the temperature and the thermal forces*, so that

(11) $$\Phi = \varphi\,(T, U_1, U_2, \dots),$$

we obtain the relations

(12) $$-\frac{\partial\Phi}{\partial T} = S, \qquad \frac{\partial\Phi}{\partial U_1} = u_1, \text{ etc.}$$

It is thus a common property of free energy and thermodynamic potential that, when they are given as functions of certain variables of state, the other quantities of state and the entropy are to be obtained by partial differentiation with respect to those variables.[3] In the case of a homogeneous system, we have simply

(13) $$\Phi = E - TS + pV.$$

### § 161. Changes of State in a Homogeneous System.

In a *homogeneous system*, there is the possibility of further special cases of change of state in addition to *adiabatic* and *isothermal* ones, *viz.*, changes which take place at *constant*

[1] Of course, only in the thermodynamical sense.

[2] The most important example of the application of Helmholtz's equation, and one treated by Helmholtz himself, is afforded by the theory of *galvanic elements*.

[3] Free energy and thermodynamic potential are consequently termed *characteristic functions*—a definition introduced by *Massieu* (about 1870).

*volume* or *constant pressure.*[1] The quantity of heat required to raise the temperature of a body by one degree at constant volume or constant pressure is called its *heat capacity* (at constant volume or constant pressure). The quotient of the heat capacity and the mass is defined as the *specific heat*, and the quotient of the heat capacity and the number of gram-molecules as the *molar heat.*

The partial differential coefficient of the volume with respect to the temperature, at constant pressure and referred to unit-volume, is called the *expansion coefficient* of the body. Denoting it by $\alpha$, we have therefore by definition

$$\alpha = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_p. \tag{1}$$

(The subscript $p$, written after the bracketed derivative, indicates that the change of state is to be carried out at constant pressure.)

The partial differential coefficient of the pressure with respect to the temperature, at constant volume and referred to unit-pressure, is known as the *pressure coefficient ;* it is thus defined by the equation

$$\beta = \frac{1}{p}\left(\frac{\partial p}{\partial T}\right)_V. \tag{2}$$

The *negative* partial derivative of the volume with respect to the pressure, at constant temperature and referred to unit-volume, is called the *compressibility coefficient*, and is hence defined by the relation

$$\kappa = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T. \tag{3}$$

Between the three quantities $\alpha$, $\beta$, and $\kappa$, there exists a simple relation which may easily be derived. Thus, if we take volume and temperature as the independent variables of state, differentiation of the equation of state gives

$$dp = \left(\frac{\partial p}{\partial V}\right)_T dV + \left(\frac{\partial p}{\partial T}\right)_V dT, \tag{4}$$

or, by eqns. (2) and (3)

$$dp = -\frac{1}{\kappa V} + \beta\, p\, dT.$$

[1] Equations describing changes of state at constant volume are also termed *isochores* or *isometrics ;* those describing changes at constant pressure are called *isobars.*

In the case of a change of state at constant pressure, for which $dp$ thus vanishes, we must therefore have

$$\left(\frac{\partial V}{\partial T}\right)_p = \beta \kappa p V.$$

Hence, by eqn. (1),

$$\alpha = \beta \kappa p. \tag{5}$$

The expansion coefficient is equal to the product of the pressure and compressibility coefficients and the pressure. If two of the coefficients be known, the third therefore follows at once.[2]

In consequence of the existence of the equation of state, we can express the *internal energy* either as a function of the temperature and the volume, or as a function of the temperature and the pressure (the possible representation of the energy as a function of the volume and the pressure does not concern us). Inasmuch as we are dealing with homogeneous bodies, in the following considerations we shall refer the quantities which, in a statistical case, are proportional to the number of individuals, to the *unit of mass*, and denote these "*specific*" quantities by the corresponding small letters. Thus, $e$ will denote specific energy, $s$ specific entropy, $v$ specific volume, and $dq$ the heat introduced into the unit of mass. Specific heat is denoted by $c$, the subscripts $v$ or $p$ being added according as it is a question of specific heat at constant volume or at constant pressure.

In accordance with the previous paragraph, we may thus put either

$$e = \varphi(T, v), \tag{6}$$

so that

$$de = \frac{\partial \varphi}{\partial T} dT + \frac{\partial \varphi}{\partial v} dv, \tag{7}$$

or

$$e = \chi(T, p), \tag{8}$$

so that

$$de = \frac{\partial \chi}{\partial T} dT + \frac{\partial \chi}{\partial p} dp. \tag{9}$$

[2] For example, the expansion coefficient of mercury at 0° C. and atmospheric pressure amounts to 0·00018. The compressibility coefficient at atmospheric pressure is 0·0000039; *i.e.*, on an increase in the pressure by one atmosphere, the unit of volume decreases by that amount. It follows from eqn. (5) that the value of the pressure coefficient ($\beta = \alpha/\kappa$) is 0·00018/0·0000039 = 46; *i.e.*, in order to maintain the volume of mercury constant while the temperature is raised 1°, a pressure of 46 atmospheres must be exerted. The accuracy of eqn. (5) is, in fact, confirmed by experiment.

Now, it follows from the *First Principal Law* that

$$dq = de + p\,dv. \tag{10}$$

Hence, if temperature and volume be taken as the independent variables of state, the First Principal Law becomes

$$dq = \frac{\partial \varphi}{\partial T}\,dT + \left(\frac{\partial \varphi}{\partial v} + p\right) dv\,. \tag{11}$$

If, on the other hand, pressure and temperature be chosen as independent variables, we have from the equation of state

$$dv = \left(\frac{\partial v}{\partial T}\right)_p dT + \left(\frac{\partial v}{\partial p}\right)_T dp\,, \tag{12}$$

and therefore, by eqns. (9) and (10),

$$dq = \left(\frac{\partial \chi}{\partial T} + p\,\frac{\partial v}{\partial T}\right) dT + \left(\frac{\partial \chi}{\partial p} + p\,\frac{\partial v}{\partial p}\right) dp\,. \tag{13}$$

Hence, we find for the specific heats

$$c_v = \frac{\partial \varphi}{\partial T}, \tag{14}$$

and

$$c_p = \frac{\partial \chi}{\partial T} + p\,\frac{\partial v}{\partial T}\,. \tag{15}$$

Putting, for brevity,

$$\frac{\partial \varphi}{dv} + p = \zeta\,, \tag{16}$$

eqn. (7) becomes, when eqn. (14) is taken into account,

$$de = c_v\,dT + (\zeta - p)\,dv\,. \tag{17}$$

Since $e$ is completely determined by $T$ and $v$, we must have

$$\frac{\partial}{\partial v}\left(\frac{\partial e}{\partial T}\right)_v = \frac{\partial}{\partial T}\left(\frac{\partial e}{\partial v}\right)_T, \tag{18}$$

the partial differentiations being carried out with the help of eqn. (17). Thus we find

$$\frac{\partial c_v}{\partial v} = \frac{\partial \zeta}{\partial T} - \frac{\partial p}{\partial T}, \tag{19}$$

where $p$ is to be regarded as a function of $v$ and $T$, so that it has to be differentiated with respect to these two quantities as independent variables.

Like the internal energy, the *entropy* can also be represented as a complete *function of volume and temperature*. According to the Second Principal Law, when the quantities

are referred to unit-mass we have for the differential of the entropy

$$ds = \frac{dq}{T}. \tag{20}$$

Hence, in accordance with eqns. (11), (14), and (16),

$$ds = \frac{c_v}{T}\,dT + \frac{\zeta}{T}\,dv. \tag{21}$$

Here, again, the relation

$$\frac{\partial}{\partial v}\left(\frac{\partial s}{\partial T}\right)_v = \frac{\partial}{\partial T}\left(\frac{\partial s}{\partial v}\right)_T \tag{22}$$

must hold good, or, by eqn. (21),

$$\frac{1}{T}\,\frac{\partial c_v}{\partial v} = -\frac{\zeta}{T^2} + \frac{1}{T}\,\frac{\partial \zeta}{\partial T}. \tag{23}$$

On multiplying eqn. (23) by $T$ and subtracting it from eqn. (19), we find

$$\zeta = T\frac{\partial p}{\partial T}, \tag{24}$$

or, bearing in mind eqn. (2),

$$\zeta = \beta\, p\, T. \tag{25}$$

Introducing this value for $\zeta$ into eqn. (11), we obtain finally, when eqn. (14) is taken into account, the following important relation which is a statement of the *First Principal Law* when *temperature* and *specific volume* are chosen as *independent variables of state*:—

$$dq = c_v\,dT + \beta\, p\, T\, dv. \tag{26}$$

In applying this equation to an *isothermal change of volume*, we have to put $dT$ equal to zero. Eqn. (26) then shows that, on an isothermal increase in volume, a body absorbs or emits heat according as its *pressure coefficient* at the temperature in question is positive or negative. If we wish to apply eqn. (26) to an *adiabatic* change of state, $dq$ must be equated to zero. An adiabatic increase in volume therefore always produces a cooling effect, provided the pressure coefficient be positive. Since this is usually the case, bodies are in general *heated by adiabatic compression*.

Again, let us investigate the form assumed by the First Principal Law when temperature and pressure are chosen as independent variables. We put, for brevity,

$$\frac{\partial \chi}{\partial p} + p\,\frac{\partial v}{\partial p} = \eta. \tag{27}$$

Eqn. (9) then becomes, when eqn. (15) is taken into account,

$$de = \left(c_p - p\frac{\partial v}{\partial T}\right) dT + \left(\eta - p\frac{\partial v}{\partial p}\right) dp . \tag{28}$$

Since $e$ can be represented as a function of the two independent variables, $T$ and $p$, we must therefore have

$$\frac{\partial}{\partial p}\left(c_p - p\frac{\partial v}{\partial T}\right) = \frac{\partial}{\partial T}\left(\eta - p\frac{\partial v}{\partial p}\right),$$

or,

$$\frac{\partial c_p}{\partial p} - \frac{\partial v}{\partial T} = \frac{\partial \eta}{\partial T} . \tag{29}$$

Moreover, we can make use of the fact that the entropy is also completely determined by $T$ and $p$. Dividing eqn. (13) by $T$, and taking eqns. (15) and (27) into account, we find

$$ds = \frac{c_p}{T} dT + \frac{\eta}{T} dp , \tag{30}$$

whence

$$\frac{\partial}{\partial p}\left(\frac{c_p}{T}\right) = \frac{\partial}{\partial T}\left(\frac{\eta}{T}\right) ,$$

or,

$$\frac{1}{T}\frac{\partial c_p}{\partial T} = -\frac{\eta}{T^2} + \frac{1}{T}\frac{\partial \eta}{\partial T} . \tag{31}$$

On multiplying this equation by $T$ and subtracting eqn. (29) from it, we obtain

$$\eta = -T\frac{\partial v}{\partial T} , \tag{32}$$

or, in view of eqn. (1),

$$\eta = -\alpha v T. \tag{33}$$

If we insert this value in eqn. (13), taking eqns. (15) and (27) into account, we obtain finally the following relation as an expression of the *First Principal Law* when *temperature* and *pressure* are chosen as *independent variables of state* :—

$$dq = c_p dT - \alpha v T dp. \tag{34}$$

In applying this equation to an isothermal process, $dT$ is again to be equated to zero. Hence, in an *isothermal increase of pressure*, a body absorbs or emits heat according as its *expansion coefficient* at the temperature in question is negative or positive. In an adiabatic change of state ($dq = 0$), the temperature rises on an increase of pressure when the expansion coefficient is positive ; when it is negative, an increase in pressure is accompanied by a fall in

temperature. In the case of *water*, the expansion coefficient is positive above 4°, but negative between 0° and 4°. Hence, if at a temperature above 4° the pressure to which water is subjected be raised, the temperature of the water rises. If the pressure be raised at a temperature below 4°, the water is cooled.

In eqn. (34) the quantity $p$ may also be negative. In that case we are dealing, not with a pressure, but with a *tension*. An example is afforded by the *extension of a rod by means of a suspended weight*, and an increase in the tension then means that $dp$ is negative. In an adiabatic change of state, an increase in the tension thus produces a cooling effect when the expansion coefficient is positive, and a heating effect when that coefficient is negative. In the case of *rubber* $\alpha$ is negative, and a rubber cord stretched by a weight becomes, in fact, warmer when the tension is increased.

With the help of the formulæ previously derived, we can easily obtain a relation connecting the specific heat at constant pressure with that at constant volume. In doing this we commence with eqn. (12). Applying eqns. (1) and (3), eqn. (12) becomes

$$dv = \alpha\, v\, dT - \kappa\, v\, dp\,. \tag{35}$$

Introducing this value into eqn. (26), we find

$$dq = (c_v + \alpha\,\beta\,p\,v\,T)\,dT - \kappa\,\beta\,p\,v\,T\,dp\,. \tag{36}$$

In accordance with eqn. (5), we may also write this as

$$dq = \left(c_v + \frac{\alpha^2}{\kappa}\,v\,T\right)\,dT - \alpha\,v\,T\,dp\,. \tag{37}$$

Comparing this with eqn. (34), we obtain the relation

$$c_p - c_v = \frac{\alpha^2}{\kappa}\,v\,T\,. \tag{38}$$

$\kappa$ is always positive, when the volume of a body decreases with increasing pressure. The specific heat at constant pressure is therefore always larger than that at constant volume, or

$$c_p > c_v\,. \tag{39}$$

In solid bodies, the quotient $\alpha^2/\kappa$ is very small, so that the difference between the two specific heats of solid bodies is generally negligible.

## § 162. The Joule-Thomson Effect.

An important application of the general laws of changes of state is to be found in a famous experiment performed by *William Thomson* and *Joule*, in 1853, and the results thereby

obtained have been in many respects of great significance for theoretical physics. This experiment essentially consists in allowing a *gas* to flow over *adiabatically* from a vessel in which, at a pressure $p'$, its specific volume is $v'$ into another vessel in which, at a pressure $p''$, its specific volume is $v''$. If the most exact application possible of the principal laws of thermodynamics to this process is to be made, the passage of the gas must be so arranged that (according to § 149) *deviations from states of equilibrium are as small as possible.* This is effected in Thomson and Joule's experiment by *forcing* the gas very slowly *through a tube* in which is fitted a *plug of wool.* The difference in pressure can easily be measured, as can the specific volume, and any chance difference in the temperatures of the gas before and after the plug.[1]

If we assume the cross-section of the cylindrical tube to be $f$, then the length of tube occupied by unit-mass of the gas before it reaches the plug is $v'/f$, whilst, after passing the plug, it is $v''/f$. To a certain extent the plug divides the tube into two gas-holders I and II. When unit-mass of the gas passes out of container I, work is thereby performed over a distance $v'/f$ *in the direction* of a force $p'f$. The work performed, or gained, is thus $p'v'$. On the other hand, since the unit-mass is compressed into container II, work is expended *against* a force $p''f$ over a distance $v''/f$. The work required to transfer unit-mass from the first to the second container is therefore given by the expression $(p''v'' - p'v')$, and, according to the Principle of the Conservation of Energy, this quantity must be equal to the decrease in internal energy $(e)$ thereby experienced by the unit-mass, seeing that the transfer is an adiabatic one. We must therefore have

$$e' - e'' = p''v'' - p'v'. \tag{1}$$

If we assume that the states of the gas in the two containers differ but little from each other, we may put

$$v'' - v' = \Delta v, \qquad p'' - p' = \Delta p, \qquad e'' - e' = \Delta e.$$

Neglecting small quantities of the second order, we can write eqn. (1) in the form

$$\Delta e + v\,\Delta p + p\,\Delta v = 0. \tag{2}$$

When pressure and temperature are chosen as independent variables of state, we have, by § 161 eqns. (10) and (34),

$$\Delta e + p\,\Delta v = c_p\,\Delta T - \alpha\, v\, T\,\Delta p.$$

[1] The fundamental idea of Thomson's experiment (the adiabatic transference of a gas from one vessel to another) had already been suggested by Gay-Lussac. The experimental method was perfected, first by Joule, and then by Thomson, who carried out his later researches conjointly with Joule.

Hence, it must follow from eqn. (2) that

$$c_p \, \Delta T + v \, (1 - \alpha \, T) \, \Delta p = 0 \, ,$$

or

$$\Delta T = \frac{v \, (\alpha \, T - 1)}{c_p} \, \Delta p \, . \tag{3}$$

This equation determines the change in temperature experienced by the gas on passing through the plug of wool, where $( - \Delta \, p)$ represents the excess pressure under which the gas is forced through the tube. As we shall see later, the experimental results afford important clues to the nature of the gases used ; but, in addition to that, they also permit of the *reduction* of the scale of a thermometer to a *thermodynamic temperature scale*, such as would be obtained by means of measurements on cyclic processes.

In order to effect this reduction, let us denote the temperature as measured on a thermometer by $\theta$, and the thermodynamic temperature by $T$ ; then we know from the definition of the temperature that, at the two fundamental points (ice-point and boiling-point), $T$ and $\theta$ coincide. Otherwise in general we have, approximately,

$$\Delta T = \Delta \theta \frac{dT}{d\theta} \, . \tag{4}$$

In like manner, we distinguish by the symbols $\alpha$ and $\alpha'$ the expansion coefficients according as they refer to the thermodynamic or thermometric temperature, and, similarly, we distinguish between $c_p'$ and $c_p$. Then, according to § 161 eqn. (1),

$$\alpha' = \alpha \frac{dT}{d\theta} \, ; \qquad c_p' = c_p \frac{dT}{d\theta} \, . \tag{5}$$

Hence, if $c_p$ be brought across to the left-hand side, eqn. (3) may be written in the form

$$c_p' \frac{d\theta}{dT} \Delta \theta \frac{dT}{d\theta} = \left( v \, \alpha' \, T \frac{d\theta}{dT} - v \right) \Delta p \, .$$

On re-writing the equation with only the quotient $dT/T$ on one side, and then integrating both sides between the ice-point (the same for both scales) and the unknown thermodynamic temperature $T$ (corresponding to a given thermometric temperature $\theta$), we find

$$\log_e T - \log_e T_0 = \int_{\theta_0}^{\theta} \frac{v \, \alpha' \, d\theta}{v + c_p' \dfrac{\Delta \theta}{\Delta p}} \, . \tag{6}$$

Now all the quantities following the integration symbol can be measured in relation to the thermometric temperature. The value of the definite integral as a function of $\theta$ can hence be derived for arbitrary values of $\theta$. If $J$ be the value of the integral, we have therefore

$$\log_e T - \log_e T_0 = J . \tag{7}$$

Again, we can evaluate the definite integral with the ice-point and boiling-point as integration limits ; let the value then obtained be $J^*$. Denoting the thermodynamic absolute temperature of the boiling-point by $T_{100}$, we have thus

$$\log_e T_{100} - \log_e T_0 = J^* . \tag{8}$$

It follows therefore from eqns. (7) and (8) that

$$T = T_0 e^J , \tag{9}$$

and

$$T_{100} = T_0 e^{J^*} . \tag{10}$$

In accordance with the definition of temperature, however, the difference between $T_{100}$ and $T_0$ must be exactly 100°, so that

$$100^\circ = T_0 (e^{J^*} - 1) . \tag{11}$$

Hence

$$T = \frac{100^\circ \cdot e^J}{e^{J^*} - 1} . \tag{12}$$

With the aid of this relation, the scale of the hydrogen thermometer can be reduced to the thermodynamic scale. As a matter of fact, the divergences are found to be very small (except at very low temperatures). Between the ice-point and the boiling-point, they do not exceed 1/20°. In that interval the hydrogen thermometer indicates too high a temperature ; above the boiling-point, the reading is too low. At 300° above the ice-point, the temperature as given by the hydrogen thermometer is nearly 1/2° too low.

## § 163. The Coexistence of Phases.

When a substance can exist in different modifications, between which transitions are necessarily accompanied by an absorption or a liberation of energy, a distinction is made in a thermodynamical sense between the different *phases* of the substance.[1] The most important example of such phases is afforded by the *states of aggregation* of a body.

[1] Phase literally means "form of appearance."

*Allotropic modifications* of a substance, however, also represent different phases. For instance, in order to transform rhombic crystalline sulphur into the monoclinic variety, heat must be introduced.

The quantity of heat required to transform unit-mass of a substance from a phase $A$ into a phase $B$ is called the *specific heat of reaction* of this transformation. When the heat of reaction is positive, $B$ is called the *higher* and $A$ the *lower phase*, and conversely when the heat is negative. Of the states of aggregation the highest is represented by the gaseous phase, and the lowest by the solid phase. The specific heats of reaction accompanying changes of the state of aggregation are known as the *heats of fusion*, *vaporization*, and *sublimation*.

We define as a *saturated complex of coexisting phases* a system which consists of different phases of a substance in thermodynamic equilibrium, its entropy thus representing a maximum for given values of the total mass, energy, and volume. The condition for the coexistence of phases in thermodynamic equilibrium is therefore

$$\delta S = 0, \tag{1}$$

with the subordinate conditions

$$\delta m = 0, \qquad \delta E = 0, \qquad \delta V = 0. \tag{2}$$

Let us consider, as a simple example, a system consisting of only *two* coexisting phases. We shall denote by $s$, $v$, and $e$ respectively the values of the entropy, the volume, and the internal energy referred to unit-mass. If the total mass of the system be $m$, let $m\,x$ be the mass of the higher phase and $m\,(1-x)$ that of the lower. We shall distinguish the quantities determining the higher phase by one dash and those determining the lower phase by two dashes, so that we have the relations.

$$s = x\,s' + (1-x)\,s'', \tag{3}$$

$$v = x\,v' + (1-x)\,v'', \tag{4}$$

$$e = x\,e' + (1-x)\,e''. \tag{5}$$

Hence, since the total mass is constant, it follows from eqn. (1) that

$$x\,\delta s' + (1-x)\,\delta s'' + (s' - s'')\,\delta x = 0. \tag{6}$$

On the other hand, in accordance with the Second Principal Law,

$$\delta s' = \frac{\delta e' + p'\,\delta v'}{T'}, \tag{7}$$

and, likewise,

$$\delta s'' = \frac{\delta e'' + p''\,\delta v''}{T''}. \tag{8}$$

We can, however, express $\delta e''$ and $\delta v''$ in terms of $\delta e'$ and $\delta v'$. For, on varying eqn. (4), we find, inasmuch as $\delta V$ must vanish [according to eqns. (2)] and the total mass is constant, that

$$x\,\delta v' + (1-x)\,\delta v'' + (v' - v'')\,\delta x = 0. \tag{9}$$

Similarly, since $\delta E$ vanishes according to eqns. (2), variation of eqn. (5) gives

$$x\,\delta e' + (1-x)\,\delta e'' + (e' - e'')\,\delta x = 0. \tag{10}$$

Solving eqns. (9) and (10) with respect to $(1-x)\,\delta v''$ and $(1-x)\,\delta e''$, and inserting the values thus obtained in eqn. (8), we find

$$(1-x)\,\delta s'' = -\frac{x\,\delta e'}{T''} - \frac{x\,p''\,\delta v'}{T''} - \left[\frac{e' - e''}{T''} + \frac{p''\,(v' - v'')}{T''}\right]\delta x.$$

Finally, introducing this relation into eqn. (6) and taking eqn. (7) into account, we obtain

$$\begin{aligned} &x\,\delta e'\left[\frac{1}{T'} - \frac{1}{T''}\right] + x\,\delta v'\left[\frac{p'}{T'} - \frac{p''}{T''}\right] + \delta x\left[s' - s''\right. \\ &\qquad \left. - \frac{e' - e''}{T''} - \frac{p''\,(v' - v'')}{T''}\right] = 0. \end{aligned} \tag{11}$$

Since the variations are arbitrary and mutually independent, eqn. (11) can only be satisfied when each of the three expressions in square brackets vanishes in itself. We must therefore have

$$T' = T'', \qquad p' = p'', \tag{12}$$

and, furthermore, if the common temperature be denoted by $T$ and the common pressure by $p$,

$$-s' + \frac{e'}{T} + \frac{v'p}{T} = -s'' + \frac{e''}{T} + \frac{v''p}{T}. \tag{13}$$

But, according to §160 eqn. (13), the thermodynamic potential referred to unit-mass is given by

$$\varphi = e - Ts + v\,p. \tag{14}$$

Eqn. (13) may hence also be written in the form

$$\varphi' = \varphi''. \tag{15}$$

*Two coexisting phases* can thus be in thermodynamic equilibrium only when both phases are at the *same pressure* and the

*same temperature*, and when unit-masses of the phases have the *same thermodynamic potential*.

It follows from the definition of thermodynamic potential (§ 160) that $\varphi'$ is a quite definite function of $p'$ and $T'$, and similarly that $\varphi''$ is a quite definite function of $p''$ and $T''$. Since, however, in accordance with eqn. (12) $p'$ is equal to $p''$ and $T'$ to $T''$, there must exist a perfectly definite connection between $p$ and $T$ in the form of an equation

$$p = \chi(T). \tag{16}$$

In the case of a saturated complex of two phases, every pressure is therefore associated with a quite definite *saturation temperature*, and, conversely, every temperature is associated with a quite definite *saturation pressure*. If we construct in a $p$-$T$-plane the curve corresponding to eqn. (16), we obtain the so-called *saturation curve*.

Since a transition between two phases is only possible as long as both phases are coexisting, a definite change of phase at a given pressure can only occur at a quite definite temperature (provided, indeed, that the transition can occur at all at that pressure). At a temperature which, at the given pressure, is lower than the corresponding saturation temperature, a substance can only exist in the lower phase. If the temperature be raised at constant pressure to the saturation temperature, the transformation of the substance from the lower to the higher phase commences. On a further addition of heat, this transformation proceeds *isothermally*. It is not until the substance has been completely transformed into the higher phase that (at constant pressure) a further rise in temperature becomes possible. It is at once evident why, at normal atmospheric pressure, every substance has a definite melting point, and why, during fusion, the temperature remains constant.[2]




FIG. 127.

If we proceed parallel to the $T$-axis from a point on the saturation curve (Fig. 127) in the direction of increasing temperature, we enter the field of the higher phase. On the other hand, if we proceed from that point parallel to the $p$-axis in the direction of increasing pres-

[2] The constancy of temperature during fusion was first discovered, simultaneously, by *Deluc* and *Black* about 1750.

sure, we must accordingly enter the field of the lower phase. If, at a given temperature, the pressure be less than the corresponding saturation pressure, the substance can exist only in the higher phase. When the pressure is isothermally increased to the saturation pressure, a transformation of the substance into the lower phase at constant pressure commences. It is only when this transformation is completed that a further increase in pressure is possible at constant temperature. The saturation pressure thus represents the *maximum pressure* at which the higher phase can exist at the temperature in question.[3]

In accordance with the equation of state, the specific volume of a substance is determined by its pressure and temperature, so that, since the saturation pressure depends only on the saturation temperature, the *specific volume* of a phase in the saturated state can be a *function only of the saturation temperature*.[4] Since the mass contained in unit-volume is reciprocal to the specific volume, a cubic centimetre of air will always contain the same mass of water vapour when saturated at a given temperature: this quantity is called the *absolute humidity* at the temperature in question. In all vapours an increase in pressure is accompanied by an increase in density, and since the saturation pressure of water vapour is its maximum pressure, it follows that the saturation density of the vapour also represents its *maximum density* for the temperature concerned. The ratio between the density of unsaturated water vapour and the maximum density at that temperature is called the *relative humidity*.[5] The temperature which would correspond to the saturation temperature for the actual density of water vapour present is known as the *dew-point*.

Let us consider the transformation of unit-mass from the lower to the higher phase at constant pressure; then, for the heat of reaction (in this case positive), the First Principal Law gives the relation

$$\lambda = e' - e'' + p\,(v' - v'')\,; \tag{17}$$

whence follows, according to eqns. (14) and (15), the simple formula

$$\lambda = T\,(s' - s'')\,. \tag{18}$$

If we remember that the phase transformation occurs

[3] The saturation pressure of water vapour is simply termed the *vapour tension*; unsaturated vapour is called *superheated*, and approximately obeys the gas laws (§ 165).

[4] Or, what amounts to the same thing, of the saturation pressure.

[5] It is generally expressed as a percentage.

at constant temperature, we obtain, with the help of § 161 eqns. (26) and (2), and bearing in mind eqn. (16), the relation

$$\lambda = \frac{dp}{dT}\, T\,(v' - v''). \tag{19}$$

When the higher phase has the larger specific volume (or, the smaller density), $dp/dT$ is positive, and the saturation temperature therefore increases with the pressure. If, on the other hand, the higher phase has the greater density, an increase in pressure produces a lowering of the transformation temperature. Since liquid water at 0° possesses a greater density than ice, the *freezing-point* is *lowered* by an increase in pressure,[6] whereas the *boiling-point* of water is *raised* thereby.

As long as the complex of two coexisting phases is saturated, it follows from what we have previously said that the state of each phase is completely defined if, besides the saturation temperature, the quantity $x$ determining the mass ratio of the two phases be given; for pressure and specific volume are alike fixed by the saturation temperature. For changes of state wherein the complex *remains saturated*, the quantities $T$ and $x$ may thus be regarded as *independent variables of state of the complex;* the quantity $x$, in fact, also represents a parameter, a change in which involves an expenditure of work $\lambda\, dx$ in amount. If heat (of amount $dq$ for unit-mass) be introduced into the complex, the state of saturation being maintained, the temperature of the complex will, on the one hand, be raised by the amount $dT$, whilst, on the other hand, the quantity $x$ will be increased by the amount $dx$. Denoting by $\gamma'$ and $\gamma''$ the *specific heats of the two phases in the saturated state*, we obtain from the First Principal Law the relation

$$dq = [x\,\gamma' + (1 - x)\,\gamma'']\,dT + \lambda\, dx. \tag{20}$$

Since, in accordance with the Second Principal Law, $dq/T$ must be the complete differential of a function determined by the two independent variables of state, $T$ and $x$, we must have

$$\frac{\partial}{\partial x}\left[\frac{x\,\gamma' + (1-x)\,\gamma''}{T}\right] = \frac{\partial}{\partial T}\left(\frac{\lambda}{T}\right),$$

or,

$$\gamma' - \gamma'' = \frac{\partial \lambda}{\partial T} - \frac{\lambda}{T}. \tag{21}$$

[6] An increase of one atmosphere in the pressure lowers the freezing-point by 0·0075° C. At a pressure of 1,000 atms., the freezing-point is no higher than −7·5° C. On the other hand, for substances other than water, which, at the melting-point, have a greater density in the solid than in the liquid state, an increase in pressure raises the melting-point; *e.g.*, such is the case with paraffin wax.

This equation, which was first established by *Clausius* (1857), enables us, for example, to calculate the specific heat of saturated water vapour when the heat of vaporization ($\lambda$) is given as a function of the saturation temperature, and also the specific heat of water is known.

Hitherto we have considered the coexistence of only two phases. In thermodynamic equilibrium this is governed by the condition that, at equal pressure and temperature, $\varphi'$ shall be equal to $\varphi''$. If *three phases* are to *coexist*, then for equal temperature and pressure we must have

$$\varphi' = \varphi'' = \varphi''' . \tag{22}$$

Since thermodynamic potential is a function of pressure and temperature, we have two equations with two unknowns for the coexistence of three phases. The solution of these equations with respect to $p$ and $T$ yields a so-called *triple-point*, in which the three saturation curves intersect. More than three phases of a single chemical substance cannot thus coexist, because otherwise there would be more than the right number of equations to determine the unknowns $p$ and $T$. Fig. 128 represents the three saturation curves of water (fusion, vaporization, and sublimation curves). They intersect at a triple-point ($\Delta$), the pressure for which amounts to 0·006 atm., and the temperature of which is 0·0075°C.[7]

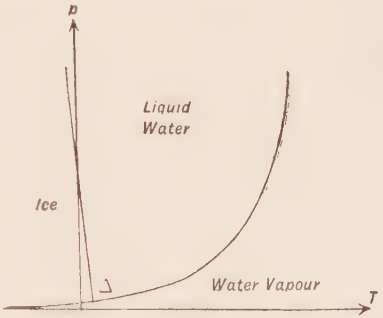


Fig. 128.

## § 164. Equilibrium in Heterogeneous Systems.

As previous considerations have shown, the thermodynamics of homogeneous systems can be referred to the relation which comprises within itself the first and second principal laws :—

$$dE = T\,dS - p\,dV . \tag{1}$$

In 1876 *Gibbs* developed this equation in an ingenious way, and, in consequence of this development, it became the basis of the *thermodynamics of heterogeneous systems*. Gibbs

[7] Since there are various allotropic modifications of ice, this triple-point is by no means the only one for water.

started with the consideration of a mixture of substances, and he regarded the *masses* of the individual components, $m_1, m_2, \ldots m_n$, as *variables*. He realized that the energy must then be not only a function of $S$ and $V$, but also a function of the individual masses, so that, in this more general case, the *complete differential of the energy* may be written in the form

$$(2)\quad dE = T\,dS - p\,dV + \mu_1 dm_1 + \mu_2 dm_2 + \ldots + \mu_n\, dm_n\,.$$

The quantities $\mu_1$, $\mu_2$, *etc.*, therefore represent the partial derivatives of the energy with respect to the individual masses at constant entropy and constant volume; for a reason which will appear later, they are now generally known as the *Gibbs potentials*.

We assume now, with Gibbs, that our system consists of *several phases*, each of which in general may contain all the $n$ substances or components of the system. We shall suppose the phases to be numbered, and the quantities associated with the first phase to be distinguished by one dash ($E'$, $S'$, $m_1'$, $m_2'$, *etc.*), those associated with the second phase by two dashes ($E''$, $S''$, $m_1''$, $m_2''$, *etc.*), and so on. On an arbitrary variation of the system, we obtain from eqn. (2), for the *variation of the energy*, the relation

$$\begin{aligned}(3)\quad \delta E = {} & T'\,\delta S' - p'\,\delta V' + \mu_1'\,\delta m_1' + \mu_2'\,\delta m_2' + \ldots + \mu_n'\delta m_n' \\ & + T''\,\delta S'' - p''\,\delta V'' + \mu_1''\delta m_1'' + \mu_2''\,\delta m_2'' + \ldots \\ & + \mu_n''\,\delta m_n'' + \ldots\end{aligned}$$

In accordance with Gibbs, we imagine our system to be contained within a rigid and non-conducting envelope, so that its total volume, its total energy, and the masses of the individual components are invariable. Analogously to § 163 eqns. (2), the following relations must then be satisfied:—

$$(4)\quad \delta E' + \delta E'' + \ldots = 0\,,$$

$$(5)\quad \delta V' + \delta V'' + \ldots = 0\,,$$

$$(6)\quad \delta m_i' + \delta m_i'' + \ldots = 0\ (i = 1 \text{ to } n)\,.$$

If, moreover, the system considered is to be in *thermodynamic equilibrium*, the variation of its total entropy must vanish; hence, in analogy to § 163 eqn. (1), we must have

$$(7)\quad \delta S' + \delta S'' + \ldots = 0\,.$$

On introducing into eqn. (4) the value given by eqn. (3) for $\delta E$, we see that the equation so obtained, together with eqns. (5), (6) and (7), can be satisfied only if

(8) $T' = T'' = \ldots, p' = p'' = \ldots,$
$\mu_1' = \mu_1'' = \ldots, \mu_2' = \mu_2'' = \ldots, etc.$

We find therefore, as a necessary condition for the *coexistence of phases* in a *heterogeneous* system, *equality of temperature and of pressure*, and also *equality of the Gibbs potentials of the individual substances for the several phases.*

In order to arrive at the physical significance of the Gibbs potentials, we shall first consider a system with a *single phase*, and integrate eqn. (2) on the assumption that the total mass of the system gradually increases from zero to a finite value in such a way that, during the increase, the proportions of the components, the specific entropy (referred to unit-mass), and the specific volume shall not vary. Since the state of the system itself, therefore, remains unchanged during the increase, temperature and pressure, and also the quantities $\mu_i$ appear as constants in the integration. Hence we obtain the relation

(9) $$E = T S - p V + \sum \mu_i m_i .$$

It follows from § 160 eqn. (13) that the *thermodynamic potential* of the system is given by

(10) $$\Phi = \sum \mu_i m_i .$$

In the special case of a single substance, we have simply $\Phi / m = \mu$, so that the *Gibbs potentials* are to be regarded as *generalizations of the thermodynamic potential referred to unit-mass.*

On differentiating eqn. (9) and subtracting eqn. (2) from the resulting relation, we obtain

(11) $$V\,dp = S\,dT + \sum m_i\,d\mu_i .$$

The *pressure* appears, therefore, to be *completely determined by the temperature and the Gibbs potentials ;* as a generalization of § 163 eqn. (16) we have

(12) $$p = \chi(T, \mu_1, \mu_2 \ldots \mu_n) .$$

It is easy also to see that, provided the *volume* be given in addition to the temperature and the potentials, the masses and the entropy are then determined ; for, by eqn. (11),

(13) $$S = V \cdot \partial p / \partial T ; \; m_i = V \cdot \partial p / \partial \mu_i .$$

If we pass on to a system with *several phases*, *each* phase is associated with an *equation* of the form of eqn. (12). When $n$ substances or components are present, an equation of this kind contains $(n + 2)$ *unknowns.* According to

eqns. (8), however, the different phases can coexist only when these *unknowns* have the *same values for the whole of the phases*. Since the number of equations cannot exceed the number of unknowns, it follows that the *number of coexisting phases* can *never* be *larger than the number of components increased by two*. Denoting the number of phases by $r$, we have

(14) $$r_{\max} = n + 2 .$$

A single substance can never exist in more than three coexisting phases, a two-component system in more than four, and so on.

When $r$ is *equal* to $(n + 2)$, the unknowns are completely determined, and the state cannot be altered without a simultaneous disturbance of the equilibrium; a system of this kind is, therefore, termed *invariant*. The state of equilibrium in question represents an $(n + 2)$-fold *multiple point*; the triple-point in a one-component system appears as a special case. If $(n + 1)$ phases coexist, the system possesses one degree of freedom, and is hence called *univariant*. When the number of phases is the same as the number of components, the system is *bivariant*; whilst if the number of phases be *less* than that of the components, the system is *multivariant*, the *number of its degrees of freedom* being determined by the relation

(15) $$f = n + 2 - r .$$

Eqns. (14) and (15) are an expression of the famous *phase rule of Gibbs*.

## § 165. The Perfect Gas.

If we consider a system of molecules, the energy of which may be regarded as purely kinetic, the theorem of equipartition leads, as we have seen in previous sections (§§ 152 and 153), to the two relations

(1) $$E = \frac{s}{2} N k T ,$$

and

(2) $$p V = N k T .$$

In these equations $N$ denotes the number of molecules, $k$ Boltzmann's constant, and $s$ the number of degrees of freedom of the molecules.

We shall now replace the number of molecules by the product of the number $(n)$ of gram-molecules and Loschmidt's

number ($L$), and we define a new *universal constant* by the relation

$$k\,L = R\,. \tag{3}$$

Eqns. (1) and (2) then become

$$E = \frac{s}{2}\,n\,R\,T\,, \tag{4}$$

and

$$p\,V = n\,R\,T\,. \tag{5}$$

Moreover, since the energy of a homogeneous body must, in general, be completely determined by volume and temperature, it follows from eqn. (4) that

$$\left(\frac{\partial E}{\partial V}\right)_T = 0\,. \tag{6}$$

In isothermal processes, the internal energy must therefore be independent of the volume. Conversely, when changes in volume are not accompanied by any change in the internal energy $E$, and thus proceed *adiabatically and without the performance of external work*, the *temperature* of such a molecular system must remain *constant*.

The properties expressed by the last three equations (4, 5, and 6) have actually been confirmed in experimental physics in the case of *gases*, the agreement between theory and experiment being the more exact the more perfect the gas, *i.e.*, the further removed from liquefaction it is under the conditions of observation. Hence, the simplest picture of a gas which we can construct is the following:—We imagine the *molecules* to be executing *rapid translatory motions*, the distances between them, however, being so large that any eventual forces acting between the molecules can be neglected, and the molecules treated *as particles*. The molecular energy is then equal to the sum of the kinetic energy of translation and the *intramolecular energy*. If we suppose (as we have already done in explaining rotational and band spectra [1]) that the molecules *rotate rigidly*, then the intramolecular potential energy does not require to be considered at all. Moreover, the internal kinetic energy must obviously be representable as a quadratic function of the time-derivatives of those coordinates which describe the internal configuration of the molecules. The hypotheses previously made therefore hold good, for the molecular energy is independent of the posi-

[1] *Vide* § 132.

tional coordinates of the molecules themselves[2]; in point of fact we have in eqn. (1), for *monatomic molecules*,

$$s = 3, \tag{7}$$

whereas, for *polyatomic gases*,

$$s > 3. \tag{8}$$

The first law to be derived empirically for gases states that, at a given constant temperature, *pressure and volume* are *inversely proportional to each other*. This law was discovered in 1660 by *Boyle*, and is therefore known as *Boyle's Law*.[3] At the beginning of the 19th century (1802), *Gay-Lussac* observed that at constant volume the pressure, and at constant pressure the volume, increases linearly with the temperature. He found also that the expansion and pressure coefficients were independent of the special nature of the gas, and identical at equal temperatures. It does actually follow from eqn. (5), and from the equations defining the expansion coefficient ($\alpha$) and the pressure coefficient ($\beta$) [§ 161 eqns. (1) and (2)], that

$$\alpha = \beta = \frac{1}{T}. \tag{9}$$

Consequently, if we observe by what fraction the volume increases at constant pressure, or the pressure increases at constant volume, when the temperature is raised from the ice-point to 1° C., this fraction is equal to the reciprocal value of the absolute temperature of the ice-point. Since a value of 1/273 was obtained for the fractional increase, the number resulting from the addition of the number 273 (more accurately, 273·1) to the measured temperature was defined as the *absolute temperature*. If the expansion coefficient be referred to the normal volume at the ice-point, instead of [as in eqn. (9)] to the volume at the temperature of observation, the expansion coefficient thus defined is equal to 1/273 at all temperatures. The same consideration applies, in corresponding fashion, to the pressure coefficient.

The universal constant $R$ is called the *gas constant*. Its value is found by measurement to be given by

$$R = 8{\cdot}315 \times 10^7 \text{ ergs/deg.}, \tag{10}$$

or, using the calorie as unit of heat,

$$R = 1{\cdot}98 \text{ cal./deg.} \tag{11}$$

[2] The coordinates meant are those determining the position of the centre of gravity of the molecules.

[3] This law is also wrongly named after Mariotte, who repeated Boyle's measurements with greater accuracy.

With the help of the previously given value for the Loschmidt number [§ 87 eqn. (6) ], we find for *Boltzmann's constant*

$$k = \frac{R}{L} = 1{\cdot}37 \times 10^{-16} \text{ ergs./deg.} \tag{12}$$

It was first experimentally discovered by *Gay-Lussac*, and subsequently confirmed by *Joule*, that scarcely any change in temperature accompanies the expansion of a gas when no external work is performed. Since, by eqn. (9), $\propto T$ does not differ from unity for an ideal gas, a temperature change could not be expected in a Joule-Thomson experiment on an ideal gas [*cf.* § 162 eqn. (3) ]. It is just refinements in the experimental methods, however, that have demonstrated the divergences of actual gases from the ideal gas laws.

A very important conclusion, to be derived directly from eqn. (2), is to the effect that, *at equal pressure and equal temperature, equal volumes of different gases* must contain the *same number of molecules*. This important law was first established in 1811 by *Avogadro*, and is therefore known by his name. It may also be stated in the form that, at given temperature and given pressure, the *densities of different gases* are *in the ratios of their molecular weights*. By means of this law, the molecular weights of various substances can be determined, and, moreover, important conclusions as to the number of atoms in molecules can be drawn. Thus, Avogadro himself discovered that one litre of hydrogen and one litre of chlorine combine to form two litres of hydrogen chloride. In accordance with Avogadro's Law, there must therefore be twice as many hydrochloric acid molecules present after the reaction as there were hydrogen molecules before it. Since, however, each hydrochloric acid molecule must contain a hydrogen atom, the hydrogen molecules must have each consisted of two atoms before the reaction. Analogous considerations apply to the chlorine molecules. In like manner, the number of atoms in the molecules of other gases and vapours could be determined. Furthermore, eqn. (2) leads directly to the so-called *Dalton's Law* (1802), which states that the pressure exerted by a *gaseous mixture* is *equal to the sum of the partial pressures* which the gases would separately exert, were each of them severally to occupy the volume actually occupied by the mixture.

In accordance with eqn. (6), the *First Principal Law* becomes, in the case of an *ideal gas*,

$$dQ = \frac{\partial E}{\partial T}\, dT + p\, dV\,. \tag{13}$$

According to the gas law expressed by eqn. (5),

$$p\, dV = d(pV) - V\, dp = n\, R\, dT = V\, dp.$$

Eqn. (13) may therefore also be written in the form

$$dQ = \left(\frac{\partial E}{\partial T} + n\, R\right) dT = V\, dp\,. \tag{14}$$

Now, for changes of state at constant volume, $dV$ is to be equated to zero in eqn. (13), whereas, for changes of state at constant pressure, $dp$ vanishes in eqn. (14). If we denote the molecular heats (referred to the gram-molecule) at constant volume and constant pressure by $C_v$ and $C_p$, we obtain from eqns. (13) and (14) the important relation

$$C_p - C_v = R\,. \tag{15}$$

For the *ratio* ($\kappa$) *of the two specific heats*[4] (whether referred to unit-mass or the gram-molecule is immaterial), we find from eqns. (13) and (14)

$$\kappa = 1 + \frac{n\, R}{\dfrac{\partial E}{\partial T}}\,,$$

or, by eqn. (4),

$$\kappa = 1 + \frac{2}{s}\,. \tag{16}$$

In the case of *monatomic gases*, $\kappa$ is hence [by eqn. (7)] equal to 5/3, while, quite generally, in accordance with eqn. (8), the relation

$$1 < \kappa \leqq \frac{5}{3} \tag{17}$$

must hold good. In point of fact, measurements made on monatomic gases have yielded the value 5/3, and the accuracy of eqn. (17) has been confirmed for polyatomic gases.

Eqn. (13) may also be written in the form

$$dQ = n\, C_v\, dT + p\, dV\,,$$

or, utilizing the gas law [eqn. (5)],

$$dQ = n\, C_v\, dT + \frac{n\, R\, T}{V}\, dV\,. \tag{18}$$

[4] The experimental determination of the specific heat at constant pressure is very much more easily effected than that at constant volume; hence the latter is indirectly calculated when the ratio $\kappa$ has been derived with the help of eqn. (21).

It follows that, in the case of an *adiabatic change of state,*

$$C_v \frac{dT}{T} + R \frac{dV}{V} = 0 . \tag{19}$$

On the other hand, by eqn. (15),

$$\frac{R}{C_v} = \kappa - 1 , \tag{20}$$

and hence, for an adiabatic change of state of an ideal gas, we find the relation

$$T V^{\kappa-1} = \text{const.} \tag{21}$$

Using the gas law to replace temperature by volume and pressure, we can write eqn. (21) in the form

$$p V^{\kappa} = \text{const.} \tag{22}$$

The *mechanical equivalent of heat* may also be easily derived from eqn. (15). For, according to the gas law, the gas constant is immediately given in absolute measure, if simply pressure and volume be observed at different temperatures. Specific heats are, however, actually measured in calories (not in ergs). If we denote the molecular heats measured in calories by $C_v'$ and $C_p'$, these quantities are equal to $C_v/J$ and $C_p/J$, where $J$ is the mechanical equivalent of heat. Hence

$$J = \frac{R}{C_p' - C_v'} = \frac{R}{C_p' \left(1 - \frac{1}{\kappa}\right)} . \tag{23}$$

By means of this relation, *Robert Mayer* was able, in 1842, to calculate the mechanical equivalent of heat from experimental data referring to the specific heats of gases.

A simple relation may also be derived from the gas laws for the *vapour pressure curve* of a liquid, *viz.*, for the saturation curve between a liquid and its vapour as represented in the $p$-$T$-plane (*cf.* § 163). Since the specific volume of the liquid is negligible in comparison with that of the vapour, § 163 eqn. (19) leads to the approximately valid equation

$$\frac{dp}{dT} = \frac{\lambda}{v T} . \tag{24}$$

According to the gas laws, however, $p\,v$ is equal to $R\,T$, so that we may also write

$$\frac{dp}{dT} = \frac{\lambda p}{R T^2} . \tag{25}$$

This formula enables us to calculate approximately the *dependence of the vapour pressure on the temperature.*

---

As we have already mentioned, the concept which refers the heat content of gases to the rapid translatory motion of molecules in all directions was developed into an exact theory by *Clausius* in 1857.[5] This theory was based on the assumption that the molecules behave as solid and completely *elastic spheres*, and execute *rectilinear* motions with *constant velocity* until they collide with other molecules or a fixed wall. The collisions will thus cause no alteration in either the total momentum or the total kinetic energy. In virtue of this concept, *gas pressure* appears to be due to the continual impacts of the molecules with the walls of the containing vessel. The hypothesis led not only to a simple derivation of the ideal gas law, but also enabled Clausius to calculate the *molecular velocities* from the experimentally determined dependency of the pressure on the temperature.[6] The values of the velocities thus found for hydrogen are about 1,700, for nitrogen about 450, and for oxygen about 420 metres per second, referred to normal atmospheric pressure and the temperature of the ice-point.

Clausius' considerations were still based on the hypothesis that, not only may the collisions between the individual molecules be neglected, but also that all the molecules in a gas possess the same velocity. Neither of these two assumptions, however, is strictly true. In 1858 *Clausius* introduced the important new concept of the *mean free path.* By this he understood the average distance traversed by a molecule between two successive collisions.[7] With the aid of a simple (but not quite accurate) consideration, he succeeded in establishing a relation between the mean free path

[5] Clausius' researches were preceded by those of Joule and Krönig by a few years. The first elements of the kinetic theory of gases, moreover, reach back to *Daniel Bernoulli*, who published in 1738 the fundamental concept of this theory in his "Hydrodynamics."

[6] If $u$ be the molecular velocity and $\varrho$ the gas density, we have simply

$$p = \frac{\varrho u^2}{3}. \tag{26}$$

[7] According to modern theory, there is, of course, no real sense in speaking of collisions. The molecules of a gas ought to be compared to miniature planetary systems, with electrical forces proceeding from their component parts. Electrical forces must therefore be acting between the molecules and constantly diverting them from their rectilinear paths, the deviations being larger the closer the mutual approach of two molecules. The molecular movements would thus be zig-zag in nature, but without any sharp corners.

and, on the one hand, the number of molecules in unit-volume and, on the other hand, the diameter of the molecules which he assumed to be spherical.[8]

In 1860, starting from quite a different point of view, *Maxwell* actually succeeded in calculating quite definite values for the mean free path. Maxwell's fundamental relation concerned the *internal friction of gases*. If, within a gas, motions occur in which numerous molecules are involved, *e.g.*, currents, it is found by experience that a force, known as internal friction, is present, this force having the direction of the velocity gradient and being proportional to it. (By velocity, of course, we mean, not the molecular, but the stream velocity.) The proportionality factor between the force referred to unit-surface and the velocity gradient is called the coefficient of internal friction. Now, from the point of view of the kinetic theory of gases, Maxwell was able to interpret internal friction as a transfer of momentum from more to less rapidly moving layers of gas. With the help of this interpretation, he obtained an equation connecting the mean free path with the coefficient of internal friction (viscosity), the density of the gas, and the molecular velocity.[9] Since the molecular velocities had already been determined by Clausius, and the coefficient of viscosity for various gases could be very accurately measured, Maxwell was able to calculate definite values for the mean free path. For example, in the case of hydrogen he found $1{\cdot}85 \times 10^{-5}$ cm.; this would mean that, on an average, a molecule experiences some ten thousand million collisions every second. This number could, of course, only have any real meaning if the distinctly arbitrary assumptions underlying the theory of gases in its usual form were actually valid.

Maxwell's paper, in which he first calculated the mean free path, has, however, been of the greatest importance for another reason. For, in this paper, he also replaced the earlier hypothesis which ascribed the same velocity to all the molecules, by his *law of distribution of molecular velocities*,

[8] If $\lambda$ be the mean free path, $n$ the number of molecules in unit-volume, and $\sigma$ the diameter of a molecule, we have

$$\lambda = \frac{3}{4\pi n \sigma^2}. \tag{27}$$

[9] If $\eta$ be the coefficient of viscosity, we have

$$\eta = \frac{\varrho u \lambda}{3}, \tag{28}$$

the other symbols being as before. Since $u$ depends only on the temperature, the viscosity is thus independent of the pressure. This result may at first appear remarkable, but it is completely confirmed by experiment.

which has been discussed in § 152. He assumed the distribution of energy among the molecules to be canonical; the concept of canonical distribution, so important for modern physics, was, indeed, suggested by Maxwell in connection with the problem of molecular velocities.[10] A uniform velocity was thus replaced by an average velocity.[11]

After Clausius and Maxwell had thus determined two of the three fundamental unknowns in the kinetic theory of gases, *viz.*, the molecular velocity and the mean free path, the most interesting problem left was the determination of the molecular diameter, *i.e.*, the evaluation of the *true size of the molecules*. This problem was first solved by *Loschmidt*, in 1865. Loschmidt started from the assumption that, when a gas is *liquefied*, the molecules occupy the whole of the volume of the liquid formed; and, on the basis of this certainly very crude hypothesis, he was able to calculate an *upper limit for the diameter of a molecule*.[12] Thus, he found for the diameter of "air molecules" (*i.e.*, of nitrogen or oxygen molecules) a value of about $10^{-7}$ cm. Now, as was mentioned above, Clausius had already derived an equation connecting mean free path and molecular diameter with the number of molecules contained in unit-volume [Footnote 8, eqn. (27)]. Hence Loschmidt was also able to determine the order of magnitude of the number of molecules contained in 1 cm.$^3$ of a gas at 0° C. and 1 atm., which number, in accordance with Avogadro's Law, must obviously be the same for all gases. He found as lower limit for this number about $10^{18}$. From the known density and molecular weight of any gas, the number of molecules contained in a gram-molecule would thus follow as about $3 \times 10^{22}$ as a lower limit, whereas the actual value of this constant, as we know now with great exactness, is about twenty times as large. After Loschmidt had thus first calculated the true size of molecules, *Boltzmann*, a year later (1866), was able to give a statistical explanation of *irreversibility* and entropy on a gas-theoretical basis by means of his so-called "*H-theorem*."[13]

[10] The designation "canonical" was, indeed, not used by Maxwell himself, but was introduced later by *Gibbs*.

[11] The ratio of the most probable to the mean velocity is $\sqrt{\pi}/2$ to 1. The mean value of the kinetic energy is 3/2 times as large as the kinetic energy corresponding to the most probable velocity.

[12] If $\gamma$ be the ratio between the volume occupied by the gas before liquefaction to that occupied after liquefaction, Loschmidt found

$$\sigma = 8\gamma\lambda. \tag{29}$$

[13] The quantity denoted by $H$ by Boltzmann is identical with the logarithm of the statistical probability, taken with the opposite sign.

Boltzmann's later investigations developed the kinetic theory of gases into a complete system.

### § 166. Osmotic Pressure.

Since the pressure of a gas is a necessary consequence of the statistical *theorem of equipartition*, as previous considerations have shown, the *molecules of any substance dissolved in a liquid* must also *tend to expand*—a tendency that manifests itself as the so-called *osmotic pressure*. Denoting this pressure by $P$, we must have the relation

$$P\,V = N\,k\,T = n\,R\,T\,, \tag{1}$$

where $R$ is again the gas constant, $N$ the number of dissolved molecules, and $n$ the number of dissolved gram-molecules.

The direct experimental detection and measurement of osmotic pressure has been rendered possible by the fact that certain substances appear *semipermeable* with respect to solutions, in that they admit of the passage of the solvent, but not of the solute. For example, fine parchment covered with *copper ferrocyanide* proves to be semipermeable as regards *aqueous solutions of sugar*; it allows the molecules of water, but not those of sugar, to pass through. Semipermeable walls play a most important part in the structure of all organisms, but a perfect artificial production of such walls has been possible only in comparatively few instances.

Just as gas pressure seeks to smooth out differences in density, so does osmotic pressure endeavour to *even out differences in concentration*. If, for example, a solution be separated by a semipermeable piston from a pure solvent (or from a less concentrated solution) contained in a cylindrical vessel, the osmotic pressure of the solution tends to drive the piston outwards, and, by admission of molecules of the solvent, to dilute the solution. If the volume of the solution confined by the piston be thereby increased by $\Delta V$, and if $\Delta n'$ gram-molecules of the solvent be thereby transferred to the solution, it follows evidently that

$$\varrho\,\Delta V = \mu\,\Delta n'\,, \tag{2}$$

where $\varrho$ is the density and $\mu$ the molecular weight of the solvent.

We shall now suppose a vessel, provided at the top with a manometer tube and closed at the bottom with a semipermeable membrane, to be filled with a solution and placed inside a larger vessel which is filled with the pure solvent. The solvent will flow into the solution, which will rise in the

manometer tube to such a height above the level in the outer vessel that the hydrostatic pressure difference is exactly equal to the osmotic pressure. In this way it appears possible to effect *measurements* of osmotic pressure; and these have actually shown that the gas law expressed in eqn. (1) is obeyed, with fair approximation, by dilute solutions. Osmotic pressure proves to be proportional both to the concentration $(n/V)$ and to the absolute temperature.

Let us next find the amount of *work* necessary to bring $\Delta n'$ gram-molecules of solvent into a solution of given osmotic pressure. Referring again to the displacement of the previously considered piston in a cylindrical vessel, we have, on the one hand, for this work

$$A = P\,\Delta V\,, \tag{3}$$

and the value of $\Delta V$ follows from eqn. (2). Hence

$$A = \frac{P\,\mu}{\varrho}\,\Delta n'\,. \tag{4}$$

On the other hand, we shall suppose a quantity of pure solvent to *evaporate* at a certain pressure $(p_0)$, and the vapour to be brought, by an *isothermal volume change*, to the vapour pressure $(p)$ of the solution. We find from the gas laws for the work performed in this isothermal process by $\Delta n'$ gram-molecules of the vapour

$$A' = \Delta n'\,R\,T\int_{V_0}^{V}\frac{dV}{V} = \Delta n'\,R\,T\log_e\frac{V}{V_0}\,. \tag{5}$$

Since, at a given temperature, the volume is inversely proportional to the pressure, we have therefore

$$A' = \Delta n'\,R\,T\log_e\frac{p_0}{p}\,. \tag{6}$$

After the pressure $p$ has been attained, let the *vapour* be *condensed in the solution*, the previously considered $\Delta n'$ gram-molecules of solvent vapour being thereby added to the solution. The work required for the whole process must, however, be equal to $A'$, because the works of evaporation and condensation mutually cancel out. Again, $A'$ must be equal to $A$, for otherwise, in contradiction of the Second Principal Law of Thermodynamics, the process of isothermal distillation considered would offer a possibility of extracting arbitrary quantities of work from the heat of solution. Therefore, equating the right-hand sides of eqns. (4) and (6), we obtain the relation

$$P = \frac{\varrho R T}{\mu} \log_e \frac{p_0}{p}. \tag{7}$$

Since, at a sufficient dilution, $p$ differs but little from $p_0$, we can use the well-known approximation formula which holds for small values of $x$:

$$\log_e (1 + x) = x .$$

We put

$$p_0 - p = \Delta p ,$$

noting from eqn. (7) that $p_0 > p$, so that the osmotic pressure leads to a *lowering of the vapour pressure*. We find, by way of approximation,

$$P = \frac{\varrho R T}{\mu} \log_e \left(1 + \frac{\Delta p}{p}\right) = \frac{\varrho R T}{\mu} \frac{\Delta p}{p}. \tag{8}$$

This relation permits of an *indirect calculation of the osmotic pressure* from observations on the lowering of the vapour pressure.

Furthermore, by combining the last equation with eqn. (1), and thus replacing $P$ by $n R T/V$, we obtain

$$\frac{\Delta p}{p} = \frac{n \mu}{\varrho V}. \tag{9}$$

But $\varrho V$ is simply the mass of the solvent, and, if we divide it by the molecular weight, we find the number ($n'$) of gram-molecules of solvent in the solution in question. Thus we derive the simple relation

$$\frac{\Delta p}{p} = \frac{n}{n'}. \tag{10}$$

The lowering of the vapour pressure stands in the same ratio to the vapour pressure itself as does the *number of (gram-) molecules of solute* to the *number of (gram-) molecules of solvent*.

Because $\Delta p$ is small, the vapour pressure curve of the solution must run approximately parallel to that of the solvent, and at a small distance from it. Let $B$ (Fig. 129) be any point on the vapour pressure curve of the solvent, and let two straight lines drawn through $B$ parallel

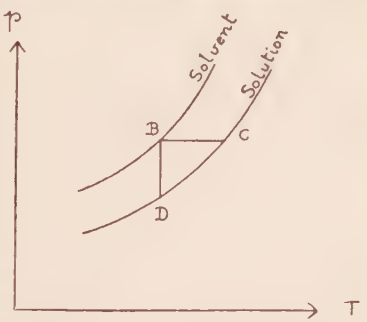



Fig. 129.

to the coordinate axes intersect the neighbouring vapour pressure curve of the solution at the points $C$ and $D$. The intercept $B\,D$ $(= \Delta p)$ denotes the lowering of the vapour pressure previously considered, whilst $B\,C$ $(= \Delta T)$ represents a *raising of the boiling-point* at the given pressure.

Since the two curves are nearly parallel and lie at a short distance from each other, we may put, by way of approximation,

$$\frac{\Delta p}{\Delta T} = \frac{dp}{dT}, \tag{11}$$

where $dp/dT$ denotes the differential coefficient of the vapour pressure curve of the pure solvent. This, however, according to § 165 eqn. (25), is approximately equal to $\lambda\, p/R\, T^2$ (where $\lambda$ is the heat of evaporation). Hence

$$\Delta p = \frac{\lambda\, p}{R\, T^2}\,\Delta T, \tag{12}$$

or, taking into account eqn. (10),

$$\Delta T = \frac{R\, T^2}{\lambda}\,\frac{n}{n'}. \tag{13}$$

Therefore the raising of the boiling-point is *proportional* to the number of gram-molecules of solute, *i.e.*, to the *concentration*. Finally, we may combine eqns. (3) and (12), obtaining for the osmotic pressure

$$P = \frac{\varrho\,\lambda}{\mu\, T}\,\Delta T. \tag{14}$$

This relation permits of an *indirect derivation of the osmotic pressure* from observations on the raising of the boiling-point of solutions.

It may be briefly mentioned that relations entirely similar to eqns. (12), (13) and (14) hold also for the *lowering of the freezing-point of solutions*, provided that heat of fusion be substituted for heat of evaporation. Historically, it may be noted that the first measurements of osmotic pressure were carried out by *Pfeffer* (1877), and that the proportionality both of the raising of the boiling-point and of the lowering of the freezing-point to the concentration was experimentally discovered by *Raoult* (1883). The identity of osmotic pressure and gas pressure was recognized in 1886 by *van't Hoff*, to whom also is due the derivation of the exact formulæ given in this section.

## § 167. The Equation of State of van der Waals.

The equation of state of the ideal gas is only approximately true for actual gases. At ordinary temperature and pressure the deviations are smallest for those gases which were formerly regarded as non-liquefiable, or "permanent." Chief among these were hydrogen, oxygen, nitrogen, and (later) the rare gases; whilst large deviations from the ideal gas law were observed, for instance, in the case of carbon dioxide. The *deviations from the ideal gas law* were first described by *van der Waals*, in 1873, in terms of an equation of state which he obtained with the help of comparatively simple theoretical considerations, and which agree fairly well with experiment. Van der Waals took into account two circumstances neglected in the derivation of the ideal gas law: first, the fact that *forces of attraction* obviously *act between the molecules* in consequence of *cohesion*,[1] and secondly, that the *volume of a gas cannot be diminished below a definite limit.*

The effect of the mutual attraction of the molecules on a gas is evidently the same as that of an external pressure. We can therefore allow for this internal attraction by supposing the *external pressure* to be *increased by an additive term.* The denser the gas, the larger must be this term; it must, indeed, evidently be proportional to the *square* of the density if, as seems quite plausible, we suppose the undetermined law of internal attraction to be similar in form to Newton's Law of Gravitation or Coulomb's Law, for in them the force varies as the *product* of two masses, or of two quantities of electricity, or of two magnetic pole strengths. Since, however, the density is inversely proportional to the specific volume ($v$), the additive term may be given the form $a/v^2$, where $a$ denotes a constant characteristic of the gas concerned.

The lowest limit to which the specific volume of a gas can ever be diminished would represent a second constant characteristic of the gas, and this is generally denoted by $b$. It seems then reasonable to replace the specific volume in the equation of state by the difference between it and this constant. The improved equation of state therefore runs

$$\left(p + \frac{a}{v^2}\right)(v - b) = RT, \tag{1}$$

[1] Debye has made an interesting attempt to calculate these cohesion forces, and thus the constant $a$ in van der Waals's equation, on the basis of the electron theory. *Vide Physikal. Zeitschr.* 21 (1920), p. 178.

for which we may also write

$$p\,v^3 - (b\,p + R\,T)\,v^2 + a\,v = a\,b. \tag{2}$$

Since this equation is of the *third degree* with respect to $v$, we recognize that, at constant temperature, each value of $p$ is associated with three values of the specific volume, of which either all three, or else only one, are real. If the state of a gas be represented graphically in a $p$-$v$-plane, the so-called *isotherms* for an ideal gas must represent *rectangular hyperbolas*, because at constant temperature the product $p\,v$ is constant. According to van der Waals's equation of state, however, these rectangular hyperbolas are replaced by *curves of the third degree* which, when the equation has three real roots, are *intersected at three points* by a straight line parallel to the $v$-axis. To a definite pressure there correspond three values $v_1$, $v_2$, $v_3$, where $v_1 < v_2 < v_3$ (Fig. 130).

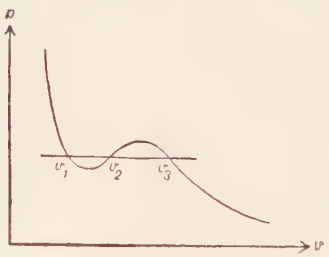


FIG. 130.

The significance of these three values can be explained, in agreement with experimental facts, by stating that, at any specific volume smaller than $v_1$, the body is liquid, but that, at any specific volume larger than $v_3$, it is gaseous, whereas, for values between $v_1$ and $v_3$, both the liquid and the gaseous states of aggregation are possible. Now the part of the isotherm lying between $v_1$ and $v_3$ shows a minimum and a maximum. It is, indeed, not possible to realize the segment of the curve lying between the minimum and the maximum, for the volume would then have to increase with increasing pressure. In certain circumstances, however, the parts of the curve extending from $v_1$ to the minimum and from the maximum to $v_3$ can be realized. The first part represents the state of *superheated liquid*; the second that of *supercooled* or, as we may say, of *supersaturated vapour*. When, as is generally the case, liquid and gaseous phases coexist in thermodynamic equilibrium, the curve is replaced by a straight line between $v_1$ and $v_3$, owing to the constancy of the pressure.[2]

If we draw a set of isotherms in accordance with van der Waals's equation (Fig. 131), we see that, with increasing temperature, the minimum and maximum of the curves draw

[2] *Vide* § 163 eqn. (12).

closer and closer together. We may thus construct a further curve, represented by a broken line in Fig. 131, which passes through the minima of all the isotherms on one branch, and through the maxima on the other branch. This curve possesses a maximum at a certain point, and the state of the gas represented by that point is termed its *critical state*. This state is defined by a *critical temperature*, a *critical pressure*, and a *critical specific volume*. At temperatures above the critical one no liquefaction at all is possible, for all higher isotherms can give only a single real root for $v$.

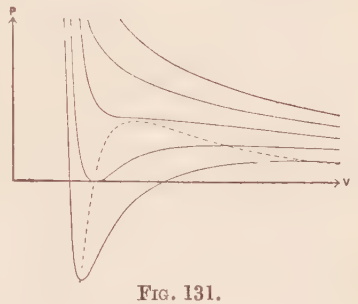


FIG. 131.

At the minimum and maximum on each isotherm the tangent to the isotherm is parallel to the $v$-axis, so that $dp/dv$ vanishes. At every point along the broken curve in Fig. 131 the relation

$$\frac{dp}{dv} = 0 \tag{3}$$

must therefore be satisfied. Now, by eqn. (1),

$$p = \frac{R\,T}{v - b} - \frac{a}{v^2}. \tag{4}$$

If we differentiate this equation with respect to $v$ ($T$ being kept constant, since we are considering only a single isotherm), and equate the differential coefficient to zero, we find in accordance with eqn. (3)

$$\frac{R\,T}{(v - b)^2} = \frac{2\,a}{v^3}. \tag{5}$$

From this, together with eqn. (4), we obtain as equation of the broken curve in Fig. 131

$$p = \frac{2\,a\,(v - b)}{v^3} - \frac{a}{v^2},$$

or,

$$p = \frac{a}{v^2} - \frac{2\,a\,b}{v^3}. \tag{6}$$

The maximum of the curve, which represents the critical state, is further characterized by the fact that the tangent to the curve is there horizontal. At the maximum, therefore, in addition to eqns. (4) and (6), which, of course, hold

good, the differential coefficient of the right-hand side of eqn. (6) with respect to $v$ must vanish. In the critical state, we must therefore have

$$\frac{d}{dv}\left(\frac{a}{v^2} - \frac{2\,a\,b}{v^3}\right) = 0\,.$$

Hence, denoting the critical quantities of state by asterisks, it follows that

$$v^* = 3\,b\,. \tag{7}$$

Introducing this value into eqn. (6), we obtain

$$p^* = \frac{a}{27\,b^2}, \tag{8}$$

and, finally, we find from eqn. (5) for the critical temperature,[3]

$$T^* = \frac{8\,a}{27\,b\,R}. \tag{9}$$

The constants $a$ and $b$, which determine the deviations from the ideal gaseous state, are smallest in the case of *helium*, and are also very small in the case of *hydrogen*, which explains why these two gases are so peculiarly suitable for thermometric purposes. The lowest critical temperature is that of helium, *viz.*, — 268° C. For hydrogen it amounts to — 240° C., for nitrogen to — 147° C., and for oxygen to — 119° C. For carbon dioxide the critical temperature is + 31° C., so that carbon dioxide can be liquefied at room temperatures. It was, indeed, with this gas that the deviation from Boyle's Law was first thoroughly investigated by *Andrews*, who thereby first arrived at the important concept of critical temperature.[4]

## § 168. Law of Mass Action.

The kinetic concept of matter permits of a simple derivation of the fundamental law of theoretical chemistry,

[3] On solving eqns. (7) and (8) with respect to $a$ and $b$, we find

$$b = \frac{v^*}{3}, \quad a = 3\,p^*\,v^{*2}, \quad R = \frac{8\,p^*\,v^*}{3\,T^*}.$$

Introducing, for the sake of brevity, the so-called *reduced quantities of state*,

$$\pi = \frac{p}{p^*}, \quad \omega = \frac{v}{v^*}, \quad \tau = \frac{T}{T^*},$$

we can write van der Waals's equation of state, as may readily be seen, in a form which is quite independent of all special properties of the gas concerned. The reduced equation of state thus becomes

$$\left(\pi + \frac{3}{\omega^2}\right)(3\,\omega - 1) = 8\,\tau\,. \tag{10}$$

[4] Half a century before Andrews' time, Cagniard de la Tour had already experimentally realized the critical state.

the so-called *law of mass action*. It is a well-known experimental fact that chemical reactions never of themselves go right through to the end, but always come to a stop when a certain *chemical equilibrium* is attained. This state of equilibrium is the same, whether reached by the direct or the reverse reaction. If, for example, a mixture of hydrogen and iodine be heated, *hydrogen iodide* is formed in accordance with the equation

$$H_2 + I_2 = 2\,HI\,. \tag{1}$$

The reaction becomes slower and slower until, at a temperature of 400° for instance, it ceases when about 79% of the hydrogen and iodine molecules have entered into combination.[1] Conversely, if hydrogen iodide be heated to 400°, *decomposition* commences: at the beginning this is rapid, but it slows up and, after some hours, ceases at a point when, again, 79% of the HI molecules remain undecomposed.

On the basis of the kinetic theory of heat, we can assume the *reactivity*, in the case of the formation of hydrogen iodide, to be *proportional to the number of collisions* between hydrogen and iodine molecules. This number, however, must in its turn be proportional, on the one hand, to the *concentration* of the hydrogen molecules, and, on the other hand, to the concentration of the iodine molecules. Concentration is usually stated in *gram-molecules per litre*, and indicated by writing the chemical symbol concerned in square brackets. The velocity of formation of hydrogen iodide may therefore be represented by the equation

$$v = k\,[H_2]\,[I_2]\,, \tag{2}$$

where $k$ is the so-called velocity constant (*i.e.*, constant for given temperature and given pressure).

Conversely, since the *decomposition* of a molecule of hydrogen iodide is to be referred to a collision between two HI molecules, the *velocity of decomposition of hydrogen iodide* may be represented by the analogous equation

$$v' = k'\,[HI]^2\,. \tag{3}$$

For the *resultant* rate of reaction, we have, therefore,

$$v^* = v - v'\,, \tag{4}$$

a positive value of $v^*$ denoting synthesis, and a negative value decomposition. Reasoning in this way, chemical

[1] It is, of course, assumed that equal numbers of hydrogen and iodine molecules were originally present: in view of Avogadro's Law, such will be the case if the initial volumes of hydrogen and iodine vapour were the same.

equilibrium appears as a *statistical equilibrium* determined by the condition

$$v = v' . \tag{5}$$

If we put, for brevity,

$$\frac{k'}{k} = K , \tag{6}$$

we obtain the condition of equilibrium in the form

$$\frac{[H_2]\,[I_2]}{[HI]^2} = K , \tag{7}$$

where $K$, the so-called *equilibrium constant*, depends, for a given reaction, only on the temperature and pressure.

Quite generally, we shall suppose a chemical reaction involving a constant number of molecules to be expressed by the equation

$$n_1 A_1 + n_2 A_2 + n_3 A_3 + \ldots \rightleftharpoons r_1 B_1 + r_2 B_2 + r_3 B_3 + \ldots \tag{8}$$

The condition of chemical equilibrium may then be represented by the equation

$$\frac{[A_1]^{n_1}\,[A_2]^{n_2}\,[A_3]^{n_3} \ldots}{[B_1]^{r_1}\,[B_2]^{r_2}\,[B_3]^{r_3} \ldots} = K . \tag{9}$$

This equation is an expression of the *law of mass action*, which was propounded in 1867 by the two Norwegian chemists, *Guldberg* and *Waage*.

### § 169. The Brownian Movement.

An average value, which is obtained from only a *small* number of individual values, may experience relatively considerable variations, not through any external cause, but simply because of *individual irregularities*. *Spontaneous fluctuations* must therefore be perceptible in all physical quantities, the value of which depends on only a small number of individual values. Conversely, the observation of such fluctuational phenomena may be regarded as a certain proof of the individual composition of matter and physical processes.

The longest known *phenomenon of molecular fluctuations* is the *Brownian movement*. This consists in a lively, irregular, *zig-zag motion* executed by very small particles suspended in a liquid or in a gas. The phenomenon was discovered in 1827 by Brown, the botanist, in connection with the floating of the pollen of plants in water, the linear

dimensions of the pollen amounting to about 1/200 mm. It was not until half a century later that *Christian Wiener* recognized the true cause of the phenomenon to lie in the internal state of motion which, in accordance with the kinetic theory of heat, must be ascribed to the liquid. An exact *theory* of the Brownian movement was developed in 1905 simultaneously by *Einstein* and by *Smoluchowski*. As this theory shows, the Brownian movement as observed does not in any way represent the *actual* motion of the suspended particle, but only the *resultant displacements* during such *intervals* as the human eye can still distinguish; the actual motion is so rapid that it escapes observation.

Let, now, $n$ be the number of particles in unit-volume; we have evidently to regard $n$ as a function of the coordinates. We choose a given point as coordinate origin, and cut off around it a region which is so small that, within it, the gradient of $n$ may be regarded as *constant* both in magnitude and direction. We shall take the direction opposite to the gradient as $x$-axis; then, if $g$ be the magnitude of the gradient $(\partial n/\partial x)$, we have

$$n = n_0 - g\,x\,. \tag{1}$$

Let us suppose that, in a given time $t^*$, $dn$ out of the $n$ particles experience a fortuitous variation of their coordinates by an amount lying between $\xi$ and $(\xi + d\xi)$. We may put

$$dn = n\,\varphi(\xi)\,d\xi\,, \tag{2}$$

where the functions $\varphi(\xi)$ must evidently satisfy the two conditions

$$\varphi(-\xi) = \varphi(\xi) \tag{3}$$

and

$$\int_{-\infty}^{+\infty} \varphi(\xi)\,d\xi = 1\,. \tag{4}$$

We shall imagine a rectangular prism, with base of unit-area, to be constructed about the $x$-axis, and to be divided into a series of slices, $dx$ in thickness, parallel to the $y$-$z$-plane. Let one of these slices be bounded by the planes with the $x$-coordinates $x$ and $(x + dx)$. Of the particles originally contained in this slice we shall segregate those for which $\xi$ lies between $\xi$ and $(\xi + d\xi)$. The number of these particles is given by the expression

$$(n_0 - g\,x)\,dx\,\varphi(\xi)\,d\xi\,.$$

On the other hand, the number of particles in the slice bounded by the planes with the corresponding *negative* values of $x$, $(-x)$ and $(-x-dx)$, is

$$(n_0 + g\,x)\,dx\,\bar{\varphi}(\xi)\,d\xi\,.$$

On subtracting the first from the second expression and integrating over $\xi$ and $x$, we obtain, with a proper determination of the integration limits, the *resultant* number of particles which, in the time $t^*$, pass across unit-area of the $y$-$z$-plane in the direction of increasing $x$. We shall denote this number by $N'\,t^*$. The integration has obviously to be taken over all possible *magnitudes* of $\xi$, or from $\xi = 0$ to $\xi = \infty$.[1] Again, for a given value of $\xi$, only those values of $x$ which are *smaller* than $\xi$, and thus lie between 0 and $\xi$, come into account; for otherwise a displacement by $\xi$ would not suffice to bring a particle into the $y$-$z$-plane.

We find therefore

$$N'\,t^* = 2\,g\int_0^\infty \varphi(\xi)\,d\xi \int_0^\xi x\,dx\,, \tag{5}$$

or, in view of eqn. (3),

$$N'\,t^* = \frac{g}{2}\int_{-\infty}^{\infty} \xi^2\,\varphi(\xi)\,d\xi\,. \tag{6}$$

Denoting by $\bar{\xi^2}$ the *mean value of the square of the displacement*, within the time $t^*$, and applying eqn. (4), we obtain the relation

$$N'\,t^* = \frac{g}{2}\,\bar{\xi^2}\,. \tag{7}$$

The mean square of the displacement therefore increases *proportionately with the time*.

It is, however, possible to derive another and quite different form of relation for the number denoted by $N'$. The totality of particles behaves as a gas (*cf.* § 170), and its *pressure* is consequently given by the formula

$$p = n\,k\,T\,, \tag{8}$$

where $n$ is the number of particles in unit-volume, and $k$ is Boltzmann's constant. According to a well-known mechanical relation—§ 47 eqn. (2)—the *force* acting on the *unit of volume* is equal to the negative *gradient of the pressure*. Dividing by $n$, we obtain the *force* $(K)$ acting on an *individual*

[1] Only *magnitudes* need be considered because the *y-z-plane* can be reached by particles with positive $x$-values only for negative values of $\xi$, and conversely.

*particle.* Since the negative gradient of $n$ is equal to $g$, we find therefore

$$K = \frac{g\,k\,T}{n}. \tag{9}$$

According to a well-known formula established by *Stokes* (*cf.* § 86), a spherical body of radius $a$ attains, in a medium with a *coefficient of viscosity* $\eta$, a constant *velocity*, the magnitude of which is given by

$$v = \frac{K}{6\,\pi\,\eta\,a}. \tag{10}$$

Experimental investigations have proved that this formula must still hold good for particles of about the size of a sugar molecule. On multiplying $v$ by the number of particles in unit-volume, we obtain the number (previously denoted by $N'$) of particles which pass across unit-area of the $y$-$z$-plane in unit-time. Hence we find from eqn. (9)

$$N' = \frac{g\,k\,T}{6\,\pi\,\eta\,a}, \tag{11}$$

or, by eqn. (7),

$$\bar{\xi}^2 = \frac{k\,T}{3\,\pi\,\eta\,a}\,t^*. \tag{12}$$

This relation, which was first derived by *Einstein*, offers the possibility, on the one hand of calculating the *radii* of the particles from experimental determinations of the mean square of the displacement, and on the other hand of *calculating the Loschmidt number*—the quotient of the gas constant and Boltzmann's constant—from observations on the Brownian movement. Both methods have yielded very satisfactory results, and it has in this way been possible to determine the Loschmidt number with an error of less than 1%.

As Einstein showed, the square of the displacement is also closely connected with the *coefficient of diffusion*. The number of particles passing across unit-area of the $y$-$z$-plane in unit-time can, as is well known, be regarded as proportional to the concentration and to the concentration gradient, *i.e.*, to the quantities $n$ and $g$. The proportionality factor, which may be denoted by $D$, represents the coefficient of diffusion. We have, therefore,

$$N' = D\,n\,g. \tag{13}$$

Hence it follows from eqn. (7) that

(14) $$\overline{\xi^2} = 2\,D\,n\,t^*\,,$$

whilst, from eqn. (12), we obtain the relation

(15) $$D = \frac{k\,T}{6\,\pi\,\eta\,a\,n}\,.$$

### § 170. Sedimentation Equilibria.

It follows from the *kinetic* concept of *osmotic pressure* that a thermal pressure must be exerted by even relatively large, and thus *individually recognizable*, *particles* which are *suspended in a liquid*. If $n$ be the concentration, or the number of particles in unit-volume, the relation

(1) $$P = n\,k\,T\,,$$

where $k$ is again Boltzmann's constant, must be satisfied. From the point of view of the kinetic theory of gases, the validity of this relation is subject only to the conditions that the linear dimensions of the suspended particles shall be small as compared with the average distances between adjacent particles, and that any forces acting between them can be neglected.

These conditions are satisfied when particles, which may be *observed with a microscope*, are suspended in a liquid, whether these particles are solid or liquid: in the former case we speak of a *suspension*, and in the latter of an *emulsion*. Frequently, the conditions stated are satisfied also with *colloidal solutions*, *i.e.*, suspensions in which the particles are too small to be observed with a microscope, but yet are large enough to appear in an *ultra-microscope* as minute plates, in consequence of their diffractional effect, although their true form cannot be recognized.

If a suspension, emulsion, or colloidal solution is contained in a vertical vessel, the dual influence of the osmotic pressure and of *gravity* must lead to a so-called *sedimentation equilibrium* of the suspended particles. Analogously to the dependence of the atmospheric pressure on the height above the earth, this equilibrium will result in a definite *vertical concentration gradient*. With the help of a simple theoretical relation, measurements of this gradient permit of a *direct determination of the Loschmidt number*.

Let us consider within the liquid a horizontal layer, with unit-area as base and with height equal to $dh$. The weight of the particles contained within this layer must then be equal to the increase in the osmotic pressure corresponding

to the height $dh$ and occasioned by the increase in concentration. Therefore, we have

$$dP = n\,m\,g\,dh\,, \tag{2}$$

where $m$ is the mass of a particle and $g$ the acceleration due to gravity. On the other hand, on differentiating eqn. (1), we find

$$dP = dn\,k\,T\,. \tag{3}$$

Combining the last two equations, and replacing $k$ by the quotient of the gas constant $R$ and the Loschmidt number $L$, we obtain

$$\frac{dn}{n} = \frac{m\,g\,L}{R\,T}\,dh\,. \tag{4}$$

An integration between two limits defined by the indices 1 and 2 gives us, therefore,[1]

$$\log_e \frac{n_2}{n_1} = \frac{m\,g\,L}{R\,T}\,(h_2 - h_1)\,. \tag{5}$$

When the mass of the individual particles is known, we can thus calculate the Loschmidt number from observations on the concentration gradient. Measurements of this kind were first carried out by *Perrin*, who, in his researches, used aqueous suspensions of particles of gamboge, with a weight of from 2 to $2{\cdot}5 \times 10^{-14}$ grams.[2] Perrin's results indicated errors of more than 10%. Subsequent determinations by *Westgren*,[3] using colloidal particles of silver and gold, yielded a value of $6{\cdot}04 \times 10^{23}$ for the Loschmidt number—a value which agrees, to within one-third of 1 *per cent.*, with that derived from electrical measurements.

For the rest, it should be noted that the number of particles contained within any delimited volume of a suspension is subject to *spontaneous fluctuations*. If $N$ be the actual number of particles at a given instant, and $N^*$ the average number of particles, the quantity

$$\epsilon = \frac{N - N^*}{N^*} \tag{6}$$

represents the average *relative fluctuation of the concentration*.

[1] It may be remarked in passing that a simple transformation of eqn. (4) leads to the well-known formula connecting the *barometric reading with the altitude*.

[2] In these experiments, the most difficult problem is to prepare particles of *equal* size.

[3] A. Westgren, *Zeitschr. f. anorg. Chem.*, **95** (1915), p. 231.

Considerations in the theory of probability, into which we cannot enter here,[4] give the relation

$$\epsilon = \sqrt{\frac{2}{\pi N^*}}. \tag{7}$$

For a volume containing about 10,000 particles, the average relative fluctuation is, therefore, approximately 1%.

Furthermore, eqn. (7) describes not only fluctuations in the concentration of suspensions, but also *fluctuations in the density of gases*. Admittedly we must choose an extremely small volume within the gas, if the fluctuations in it are to be noticeable, for, if a cube of air is to contain 10,000 molecules, the length of a side will be only $7 \times 10^{-6}$ cm. Although the fluctuations in density become thus insignificant over macroscopic regions, yet they may easily produce an *optical non-homogeneity* in a gas, or a *slight opalescence*. On traversing an opalescent medium, however, light-waves experience a *diffraction* which is the stronger, the *smaller* the *wave-length*. Hence, in the diffracted light, the colour blue will be more intense than in white sunlight. In this fact lies the explanation of the *blue colour of the sky*, as *Smoluchowski* recognized. Smoluchowski's theory even permitted of a determination, with an accuracy of about 6%, of the *Loschmidt number* from observations on the intensity ratio between the blue of the sky and sunlight.[5]

[4] *Cf.* R. Fürth, "*Schwankungserscheinungen in der Physik*" (Brunswick Sammlung Vieweg, 1920: Part 48).

[5] *Cf.* H. Dember, *Ann. der Phys.*, **49** (1916), p. 590.

## Chapter XX

# THE QUANTUM THEORY OF HEAT

### § 171. The Specific Heat of Solid Bodies.

From the statistical point of view we have seen that the heat content of gases is to be regarded as a form of translational energy, and, similarly, statistical theory refers the *heat content of solid bodies to atomic oscillations*. In doing this, use is made of the conclusion, which follows from the quantum principle, that oscillatory energy is composed of *energy-elements* which are proportional to the frequency.

The simplest possible concept of this kind is the assumption that all the atoms in a solid body oscillate with the *same frequency*, which may be denoted by $\nu^*$. The quantity $h\nu^*/k$, which must obviously have the dimensions of a temperature, may then be defined as the *characteristic temperature* ($T^*$) of the body concerned. Since the modulus of distribution of the oscillatory energy is to be equated to the heat modulus, *i.e.*, to the product of $k$ and the temperature $T$, we find as mean value for the energy of an oscillation [according to § 155 eqn. (5)]

$$\bar{\epsilon} = \frac{k\,T^*}{e^{\frac{T^*}{T}} - 1}. \tag{1}$$

If we consider an *ideal monatomic solid body*, the oscillations of each atom can be resolved into *three* mutually perpendicular and independent partial oscillations. On the above assumptions, therefore, the oscillatory energy of a gram-atom is given by

$$E = 3\,L\,\bar{\epsilon}, \tag{2}$$

where $L$ denotes the Loschmidt number. Introducing the value for $\epsilon$ following from eqn. (1), and differentiating with respect to the temperature, we obtain the specific heat referred to a gram-atom, or the so-called *atomic heat* (denoted by $C$). As a result of the differentiation we find

$$C = \frac{3\,L\,k}{\left(e^{\frac{T^*}{T}} - 1\right)^2} \cdot \left(\frac{T^*}{T}\right)^2 \cdot e^{\frac{T^*}{T}}. \tag{3}$$

This equation assumes an especially simple form when $T$ is so much larger than $T^*$ that, by way of approximation, we may put

$$e^{\frac{T^*}{T}} - 1 = \frac{T^*}{T}, \qquad e^{\frac{T^*}{T}} = 1$$

[making use of the formula according to which, for small values of $y$, $e^y$ is approximately equal to $(1 + y)$]. In that case, because the product $k\,L$ represents the gas constant,

$$C = 3\,R. \tag{4}$$

Since the value of the gas constant in thermal measure is 1·98 gram-calories [§ 165 eqn. (11)], it follows that, when the special assumptions for the validity of eqn. (4) are satisfied, the *atomic heats of all solid substances must have the same value of* 5·94 *calories*.

This law was, in fact, experimentally discovered by *Dulong* and *Petit* as long ago as 1819. The determinations then carried out showed that, for the majority of solid elements, the atomic heats at constant pressure differed by less than 10 per cent. from an average value of 6·4 cal., and a reduction of this value to constant volume [1] yields just 5·9 calories. On the other hand, no agreement was found to exist for some elements of low atomic weight, such as beryllium, boron, and especially *diamond*, the atomic heat of which at room temperature proved to be only 1·7 cal., whilst at — 50° C., as was discovered in 1875 by *H. F. Weber*, it falls to as low as 0·7 cal.

No explanation for these *deviations* from Dulong and Petit's Law could be given on the basis of "classical" statistics, the deductive processes of which, indeed, necessarily led to that law as being generally valid (we cannot enter into further details here).[2] From the point of view of the quantum theory, however, the decrease in the specific heat with the temperature immediately became intelligible, as did the fact that these deviations are particularly noticeable for elements of low atomic weight. For, if two oscillating particles are bound to their position of rest by equally large forces, it follows from the formulæ of mechanics that the frequency of the lighter particle will be greater than that of

[1] Cf. § 161 eqn. (38).

[2] Classical statistics ascribed an amount of energy $k\,T/2$ to each degree of freedom of an oscillation, and further assumed the potential energy to be equal to the kinetic energy.

the heavier,[3] and the higher frequency is associated with a higher characteristic temperature. Thus the characteristic temperature of one element may be far below room temperature, and that of another element far above ; and examples of the latter case, involving the largest deviations from Dulong and Petit's Law, may well be expected with the lightest elements.

In the light of modern atomic theory, Dulong and Petit's Law appears simply as a *special case* of a more general expression for the atomic heat, which is approximately represented by eqn. (3). This equation was derived in 1907 by *Einstein*, and formed the first application of the quantum principle to mechanical oscillations. Exact determinations carried out by *Nernst* and other investigators subsequently to the formulation of Einstein's law have shown, in fact, that the atomic heats of all substances converge at sufficiently high temperatures towards $3R$, and at sufficiently low temperatures towards zero.[4] The atomic heat of diamond ($C_v$) at about 1,000° C. amounts to 5·5 cal. On the other hand, the atomic heat of copper, for example, falls to 3·38 at 88° K.,[5] to 0·54 at 33° K., and to 0·22 cal. at 23° K. The decrease is certainly most marked with diamond. At 86° K. the atomic heat is but 0·03 cal., and at 23° K. it was found by measurement to be 0·00 cal. ; *i.e.*, at these low temperatures the idea of temperature is practically meaningless in the case of diamond.

In order to be able to test quantitatively the accuracy of Einstein's formula, it is above all necessary to determine the characteristic temperature, and thus the characteristic frequency of various substances, and, in this connection, several available paths for research have been found. In the first place, as *Lindemann* has shown, a simple relation exists between the characteristic frequency and the *temperature of the melting-point*, if we make the plausible assumption that, at the melting-point, the amplitude of the atomic oscillations becomes equal to the distances between the atoms, so that

[3] If the force attracting a particle of mass $m$ to its position of rest from a distance $x$ be $(-\gamma^2 x)$, it follows from § 6 that the frequency of the oscillation is given by

$$\nu = \frac{\gamma}{2\pi\sqrt{m}}.$$

[4] In the neighbourhood of the melting-point, the value $3R$ is generally actually exceeded, owing to the commencement of further oscillations.

[5] Absolute temperatures are generally denoted by the letter K ("Kelvin").

atomic collisions occur.[6] Furthermore, *Madelung* and then *Einstein* have demonstrated that the *elasticity constants*, particularly the *modulus of compressibility*, are determined by the characteristic frequency. Finally, *Nernst* has drawn attention to the connection existing between the characteristic frequencies and the wave-lengths (measured by *Rubens*) of the so-called *residual rays* from *diatomic crystals* such as potassium chloride.

The values of the characteristic frequencies derived in these different ways agree quite well with each other ; the values obtained are of the order of magnitude of $10^{12}$ sec.$^{-1}$, in diamond as high even as $10^{13}$ sec.$^{-1}$, so that the frequencies all lie well within the *infra-red*. They are associated with characteristic temperatures which, for example, for lead, silver, copper, and aluminium, amount to 100°–400° K., and for diamond to about 2,000° K. It was found, however, that, using these values, Einstein's formula does not reproduce the temperature course of the atomic heats accurately ; at low temperatures it invariably gives too *small* values. *Nernst* and *Lindemann* therefore formulated an *empirical* expression which represents the actual temperature course very exactly. This formula was obtained by adding to the expression on the right-hand side of Einstein's formula [eqn. (3) ] another term of precisely similar form, but with the fraction $T^*/T$ everywhere replaced by the fraction $T^*/2T$. *Einstein* himself realized that the imperfection of his formula was to be ascribed to the arbitrary assumption of a uniform frequency. A better agreement was only to be expected with a theory which took into account a *multiplicity of frequencies*, as, in a first rough approximation, was done in the Nernst-Lindemann formula, a frequency $\nu/2$ being introduced besides the principal frequency $\nu$.

From the development of this idea has arisen the theory of *atomic lattices*, which has already been briefly mentioned. On the other hand, in 1912 *Debye* developed a simple theory by *disregarding any atomistic structure of solid bodies*. He viewed them in the light of a *continuum filled with stationary elastic waves*, and applied to these waves the principle which, it is true, was first deduced by *Rayleigh* and *Jeans* (*vide* § 154) for stationary electro-magnetic waves.

[6] The formula proposed by Lindemann contains, in addition to the frequency and the temperature of the melting-point, only the atomic weight and the atomic volume ; a proportionality factor was empirically determined.

### § 172. Nernst's Heat Theorem.

Let us now consider a body filled with *stationary elastic waves* in the sense of Debye's concept; then we must obviously distinguish between *longitudinal* waves with one degree of freedom, and *transverse* waves with two degrees of freedom. We shall denote the velocity of the longitudinal waves by $v_1$, and that of the transverse waves by $v_2$, both these quantities (according to § 49) being determined by the density and the elasticity constants. The number of longitudinal vibrations, the frequencies of which lie between $\nu$ and $(\nu + d\nu)$, is given by Rayleigh and Jeans's principle [§ 154 eqn. (12)], and we have

$$dZ_1 = \frac{4\pi V \nu^2}{{v_1}^3}\, d\nu\,, \tag{1}$$

where $V$ is the volume of the body. Similarly, the number of transverse vibrations with frequencies between $\nu$ and $(\nu + d\nu)$ is given by

$$dZ_2 = \frac{4\pi V \nu^2}{{v_2}^3}\, d\nu\,. \tag{2}$$

Debye now assumes that the number of vibrations is not unlimited, but that there is a *maximum frequency* ($\nu_{\max}$) determined by the condition

$$\int_0^{\nu_{\max}} dZ_1 + 2\int_0^{\nu_{\max}} dZ_2 = 3N\,, \tag{3}$$

where $N$ is the total number of atoms in the body; for the number of degrees of freedom of each of the longitudinal vibrations is one, of each of the transverse vibrations two, and of each of the atomic vibrations three. If we put, for the sake of brevity,

$$\frac{1}{{v_1}^3} + \frac{2}{{v_2}^3} = \frac{3}{v^3}\,, \tag{4}$$

where $v$ may be called the mean wave velocity, we have

$$3N = \frac{12\pi V}{v^3}\int_0^{\nu_{\max}} \nu^2\, d\nu\,,$$

or,

$$\nu_{\max} = v\sqrt[3]{\frac{3N}{4\pi V}}\,. \tag{5}$$

As was mentioned above, the mean wave velocity can be

calculated from the density and the elasticity constants. The quotient $N/V$ represents the number of atoms in unit-volume, and we know this number from the density and the molecular weight of the substance. The maximum frequency can therefore be found from eqn. (5), and the values thus obtained agree in order of magnitude with those which had been calculated for the characteristic frequency on the basis of the more primitive assumption of a uniform frequency.

For the total *energy density* we find [according to § 155 eqn. (8) ] the value

$$\eta = \frac{12\,\pi h}{v^3} \int_0^{\nu_{\max}} \frac{\nu^3\, d\nu}{e^{\frac{h\nu}{kT}} - 1}. \tag{6}$$

Now, if the temperature be sufficiently low, $h\nu/kT$ becomes relatively large, and hence the integrand in eqn. (6) decreases very rapidly with increasing $\nu$. In this special instance of sufficiently low temperature, we may therefore replace the integration limit $\nu_{max}$ in eqn. (6) by $\infty$. If, however, we do this, the laws derived in § 155 become valid. Even if we admit the possibility of the internal energy of a body possessing a component which is independent of the temperature, the differential coefficients of the energy and the free energy with respect to the temperature must at any rate still be proportional to the cube of the temperature (according to § 155).

The differential coefficient of the energy with respect to the temperature simply represents the *specific heat*; at *low temperatures*, the specific heat of solid bodies must therefore be *proportional to the cube of the temperature*. This law was derived by *Debye* from his theory, and measurements carried out by Nernst and other investigators on diamond, copper, and other substances have actually confirmed it.

Entropy is connected with energy and free energy by the relation (frequently quoted)

$$S = \frac{E - F}{T} \tag{7}$$

[§ 160 eqn. (1) ]. The following two equations must therefore be satisfied :—

$$\lim_{T \to 0} \frac{dF}{dT} = 0\,; \qquad \lim_{T \to 0} S = 0\,. \tag{8}$$

According to these equations, both the *temperature coefficient of the free energy and the entropy vanish at the absolute zero.*

On these two relations is based an important theorem which —independently of the quantum theory—was derived by *Nernst* in 1906, and is hence generally known as *Nernst's Heat Theorem*. It is assumed to hold good not only for solid bodies, but quite generally for "*condensed*" (*i.e.*, both solid and liquid) systems.

As we have seen, entropy depends only on the temperature, and hence, at the lowest temperatures, we must have

$$\left(\frac{\partial S}{\partial p}\right)_T = 0 . \tag{9}$$

Again, multiplying § 161 eqns. (30) and (33) by the mass of the body ($m$), we find

$$dS = \frac{m c_p}{T} dT - \frac{\partial V}{\partial T} dp . \tag{10}$$

Eqn. (9) can thus only be satisfied when

$$\frac{\partial V}{\partial T} = 0 . \tag{11}$$

At very low temperatures, the volume of a solid body must therefore be independent of the temperature, so that the *expansion coefficient vanishes*. The accuracy of this theorem has actually been proved by experiments of *Grüneisen*.

### § 173. The Absolute Value of Entropy.

By the thermodynamical definition of *entropy*, only its *differential* is determined, so that every expression for entropy must necessarily contain an *additive constant* which, from the "classical" point of view, is *not determinable*. If we denote by $C_v$ the specific heat at constant volume referred to the gram-molecule, and by $n$ the number of gram-molecules, we have for the entropy differential

$$dS = \frac{dQ}{T} = \frac{n C_v dT + p dV}{T} . \tag{1}$$

In the special case of an *ideal gas*,

$$p dV = d(pV) - V dp = n R dT - n R T \frac{dp}{p} . \tag{2}$$

Finally, bearing in mind that, in accordance with § 165 eqn. (15), $(C_v + R)$ is equal to the molecular heat at constant pressure, we find

$$dS = n C_p \frac{dT}{T} - n R \frac{dp}{p} . \tag{3}$$

Hence, if we denote by $s_0$ the *entropy constant*, referred to the *gram-molecule*,

$$S = n C_p \log_e T - n R \log_e p + n s_0 . \tag{4}$$

In the special case of *monatomic gases*, modern physics has succeeded, on the one hand, in *experimentally* deriving the entropy constant with the help of *Nernst's heat theorem*, and, on the other hand, in calculating it by the means of *methods of quantum statistics*. A comparison of the two results affords a confirmation of the quantum principle.

In order to determine the entropy constant with the aid of Nernst's heat theorem, we start from a relation derived in a previous section—§ 163 eqn. (18)—

$$s' - s'' = \frac{\lambda}{T} ; \tag{5}$$

both the entropy ($s$) and the heat of evaporation ($\lambda$) are now referred to the gram-molecule, and we shall suppose $s'$ to be the entropy of the vapour and $s''$ that of the condensed phase. Furthermore, if we differentiate § 163 eqn. (17) with respect to the temperature, we find, on taking § 161 eqn. (15) into account,

$$\frac{d\lambda}{dT} = C_p' - C_p'' . \tag{6}$$

For the purpose of deriving the additive entropy constant, we may, of course, consider the system at *arbitrarily low temperatures*. According to Nernst's heat theorem, however, at very low temperatures both the entropy and the specific heat of the condensed phase are negligible in comparison with the entropy and the specific heat of the vapour, which still obeys the gas laws. With monatomic gases, to which we shall confine ourselves in the following considerations, $C_p$ is equal to $\frac{5}{2} R$, and hence [1]

$$\frac{d\lambda}{dT} = \frac{5}{2} R . \tag{7}$$

Again, denoting the heat of evaporation [2] at the absolute

[1] According to § 165 eqn. (5), $C_v$ is equal to $\frac{3}{2} R$ for monatomic gases, and it follows from § 165 eqn. (15) that $C_p$ is equal to $\frac{5}{2} R$.

[2] Instead of heat of evaporation, we should really speak more generally of the heat of reaction for the transformation from the condensed into the vapour phase.

zero by $\lambda_0$, we may put, by way of approximation, at very low temperatures

$$\lambda = \lambda_0 + \frac{d\lambda}{dT} T . \tag{8}$$

Hence, according to eqn. (5), in which we may disregard $s''$ in comparison with $s'$,

$$s' = \frac{\lambda_0}{T} + \frac{5}{2} R . \tag{9}$$

On dividing eqn. (4) by the number of gram-molecules (thus referring the entropy to the gram-molecule) and comparing eqns. (4) and (9), we obtain for the *entropy constant* referred to the gram-molecule

$$s_0 = R \log_e p - C_p \log_e T + \frac{\lambda_0}{T} + \frac{5}{2} R . \tag{10}$$

The heat of evaporation at the absolute zero ($\lambda_0$) can be determined by extrapolation, if the temperature course of $\lambda$ in the neighbourhood of the absolute zero be known. Consequently, the entropy constant appears as a constant characteristic of every substance, and independent of temperature and pressure. It is commonly replaced by another constant which, defined by the equation

$$i = \frac{s_0}{R} - \frac{5}{2} , \tag{11}$$

is known as the *chemical constant*. According to eqn. (10), this is determined (since $C_p/R$ is equal to 5/2) by the relation

$$i = \log_e p - \frac{5}{2} \log_e T + \frac{\lambda_0}{T} . \tag{12}$$

In order, now, to derive the entropy constant on the basis of *quantum-statistics*, the simplest method is to start with the *principle of Rayleigh and Jeans*, and combine it with the concept of *material waves*. We shall first effect a slight transformation in Rayleigh and Jeans's principle, as expressed in § 154 eqn. (11), by replacing the frequency by the *wave-number*, which is equal to the quotient of the frequency and the wave-velocity, and represents the number of waves in a unit of length. § 154 eqn. (11) becomes then

$$\Delta Z = 4 \pi V z^2 \Delta z , \tag{13}$$

where $V$ denotes the volume, $z$ the wave-number, and $\Delta Z$ the number of characteristic vibrations lying in the interval between $z$ and $(z + \Delta z)$. This formula, which must hold

for waves of all kinds, may now be applied to the *material waves* of a monatomic *gas* enclosed in a *hollow space*.

According to the fundamental relation of the theory of material waves [§ 91 eqn. (2)], we have for the *momentum* of a molecule

$$p = z\,h\,. \tag{14}$$

We find, therefore,

$$\Delta Z = \frac{4\,\pi\,V}{h^3}\,p^2\,\Delta p\,, \tag{15}$$

where $\Delta Z$ represents the number of possible values for the momentum in the interval between $p$ and $(p + \Delta p)$.

We shall now suppose a "*momentum space*" to be associated with the gas under consideration in such a way that every molecule is represented by a momentum point, the coordinates of which are identical with the momentum components of the molecule. The distance of the momentum point from the origin of coordinates is thus equal to $p$. If, therefore, we construct in the momentum space two concentric spherical surfaces, with radii $p$ and $(p + \Delta p)$, about the origin, the spherical shell enclosed by these two surfaces will contain a total of $4\,\pi\,p^2\,\Delta p$ momentum points, with magnitudes lying between $p$ and $(p + \Delta p)$.

It appears from eqn. (15), however, that there cannot be more than $\Delta Z$ different values of the momentum in this partial zone. We may therefore suppose this shell-like zone of the momentum space to be divided up into $\Delta Z$ *cells* in such a way that, within a cell, no distinction can be made between different values of the momentum, so that only a single momentum value is possible for each cell. We obtain the *size of a cell*, therefore, on dividing the volume of the spherical shell, $4\,\pi\,p^2\,\Delta p$, by $\Delta Z$ :—

$$\sigma = \frac{h^3}{V}\,. \tag{16}$$

In our earlier considerations of the theorem of equipartition (§ 152), we imagined the momentum space to be divided into *partial zones*, the size of which ($\Delta\tau$) was left undetermined : eqn. (15) provides us now with the definite value of $h^3/V$ for $\Delta\tau$. Multiplying § 152 eqn. (13) by $k$, let us replace $k \log_e W$ by the entropy $S$, and the modulus of distribution $\Theta$ by $k\,T$. Finally, we have to remember that, at the absolute zero, the gas is *condensed*, so that we have to subtract from the expression that we find for the entropy of the gas an amount $S^*$, which represents the entropy of the

condensed gas at the absolute zero. In this way we obtain

$$(17)\quad S = N k \left[\frac{3}{2}\log_e T + \log_e V + \log_e \frac{(2\pi e m k)^{\frac{3}{2}}}{h^3}\right] - S^* .$$

We can eliminate the volume from this formula if, in accordance with the gas law, we put

$$(18)\quad \log_e V = \log_e N + \log_e k + \log_e T - \log_e p .$$

Furthermore, as regards the *entropy of the condensed gas at the absolute zero*, we may plausibly assume that the corresponding statistical probability is represented by the number of permutations of the $N$ molecules,[3] so that

$$(19)\quad S^* = k \log_e (N\,!\,) .$$

This equation may be transformed with the help of the approximation formula previously used—§ 146 eqn. (9)—

$$(20)\quad S^* = N k (\log_e N - 1) .$$

On introducing the values given by eqns. (18) and (20) into eqn. (17), replacing $N k$ by $n R$, and remembering that $\log_e (e^{\frac{3}{2}})$ is equal to $\frac{3}{2}$, we find

$$(21)\quad S = n R \left[\frac{5}{2}\log_e T - \log_e p + \log_e \frac{(2\pi m)^{\frac{3}{2}} k^{\frac{5}{2}}}{h^3} + \frac{5}{2}\right] .$$

Since the molecular heat at constant pressure is, as already mentioned, equal to $\frac{5}{2} R$, a comparison of eqn. (21) with eqns. (4) and (11) yields, as the value of the *chemical constant*,

$$(22)\quad i = \log_e \frac{(2\pi m)^{\frac{3}{2}} k^{\frac{5}{2}}}{h^3} .$$

If we replace the molecular mass $m$ by the quotient of the *molecular weight* ($\mu$) and the Loschmidt number ($L$), we may also write eqn. (22) in the form

$$(23)\quad i = \log_e \frac{(2\pi)^{\frac{3}{2}} k^{\frac{5}{2}}}{h^3 L^{\frac{3}{2}}} + \frac{3}{2}\log_e \mu ,$$

or, substituting for the universal constants their known values,

$$(24)\quad i = 10{\cdot}17 + 1{\cdot}5 \log_e \mu .$$

With monatomic argon, with the monatomic vapours of

[3] For the quantum-statistical basis of this assumption, *vide*, *e.g.*, E. Schrödinger, *Physikal. Zeitschr.*, **25** (1924), p. 41.

mercury, zinc, and cadmium, and with hydrogen, which behaves thermally at very low temperatures like a monatomic gas, experimental measurements have, in fact, yielded for the chemical constant values which agree, within an error of a few *per cent.*, with the values calculated from the quantum-theoretical formula (24).

### § 174. Cavity Radiation.

Stationary vibrations may be of an electromagnetic as well as of a mechanico-elastic nature. In that case, however, the number of degrees of freedom will only be two, instead of three, because electromagnetic waves are always purely transversal, longitudinal waves being impossible. A *system of stationary electromagnetic waves* filling an enclosed space is defined as *cavity radiation*. The theorems derived in § 155 will, of course, be applicable to a system of this kind, especially as there is no question of any limiting maximum frequency, which only had to be taken into account in considering atomic oscillations.

Now *electromagnetic energy* can be *absorbed* and *emitted by atoms*. An *interchange of energy* between the cavity radiation and a body composed of molecules (which may be either a portion of the boundary or exist within the enclosure) is therefore possible. When, despite the possibility of an energy interchange, the energy values neither of the cavity radiation nor of the body alter, *thermodynamic equilibrium* is said to exist. In accordance with the general principles of statistics (§ 149), this is only possible when the modulus of distribution of the electromagnetic energy is equal to the heat modulus of the body; the modulus of the cavity radiation must thus be equal to the product of Boltzmann's constant $k$ and the temperature of the body with which the radiation is in thermodynamic equilibrium. This temperature may be defined as the *temperature of the cavity radiation*. Since the suppositions of § 155 hold good here, the energy density can depend only on the modulus of distribution, being proportional to its fourth power. In *thermodynamic equilibrium*, therefore, the *radiation density* can be a *function only of the temperature*, and must be *proportional to the fourth power of the temperature*.

Again, we must regard the equilibrium of a body which is *both radiating and receiving radiation* in accordance with a concept developed in 1809 by *Prévost, viz.*, as consisting in the *emission and absorption of equal amounts of energy* in any

given time.[1] The amount of energy radiated by unit-surface of the body in unit-time is known as the *emissive power* of the body. Denoting it by $K'$, we find for the quantity of energy emitted by a surface-element $df$ in an interval of time $dt$

$$dE = K' \, df \, dt. \tag{1}$$

Again, the total quantity of energy *incident* on the surface-element $df$ in the time $dt$ is given by

$$dE' = f(T) \, . \, df \, dt, \tag{2}$$

for the radiation density depends only on the temperature, so that the quantity by which $df \, . \, dt$ is multiplied on the right-hand side of the equation can be a function only of the temperature. Of the incident energy a fraction $\varrho \, . \, dE'$ will, in general, be reflected. The quantity $(\gamma = 1 - \varrho)$ is called the *absorption coefficient*, for the quantity of energy $\gamma \, dE'$ is absorbed by the irradiated body. In thermodynamic equilibrium we must have

$$dE = \gamma \, dE', \tag{3}$$

and hence, by eqns. (1) and (2),

$$\frac{K'}{\gamma} = f(T). \tag{4}$$

On the supposition of thermodynamic equilibrium, therefore, the *ratio of the emissive power to the absorption coefficient* at a given temperature must be the *same for all bodies*. This important theorem was first propounded in 1859 by *Kirchhoff*. If, in particular, $\gamma$ is equal to unity, so that the body has no reflective power at all, the body is said to be *perfectly black*.[2] The *emissive power of a perfectly black body* therefore *depends only on the temperature*. Again, it follows from eqns. (2) and (4) that the quantity of energy incident in unit-time on unit-surface of an arbitrary body is, in thermodynamic equilibrium, equal to the emissive power of a black body at the same temperature, which may be denoted simply by $K$.

Emissive power and radiation density are connected by a simple relation, which may easily be established. The amount of energy incident on unit-area of the enclosing surface in unit-time is equal to $\frac{1}{2} \eta v^*$, where $v^*$ is the mean value of the velocity component of the incident light-quanta

[1] Prévost, it is true, regarded heat as a material substance.

[2] The nearest approach to a perfectly black body among actual substances is soot. When $\gamma$ is equal to zero, the body is said to be perfectly reflecting.

with respect to the normal to the surface. The factor $\frac{1}{2}$ has to be inserted because, of the two opposed possibilities of a motion towards the surface and of a motion from it, only the former is to be considered.

Thus the quantity of energy incident in unit-time is given by the expression $\frac{1}{2}\eta c\,\overline{\cos\theta}$, the mean value of $\cos\theta$ being taken over *half a sphere* of unit-radius. According to a well-known elementary formula, an elementary cone cuts out of the surface of this half sphere an area which is given by

$$d\omega = \sin\theta\, d\theta\, d\varphi\,, \tag{5}$$

where $\theta$ can assume all values from 0 to $\pi/2$, and $\varphi$ all values from 0 to $2\pi$. We find therefore

$$\overline{\cos\theta} = \frac{1}{2\pi}\int_0^{\frac{\pi}{2}}\int_0^{2\pi}\cos\theta\sin\theta\, d\theta\, d\varphi = \int_0^{\frac{\pi}{2}}\sin\theta\, d(\sin\theta) \tag{6}$$

$$= \left|\frac{1}{2}\sin^2\theta\right|_0^{\frac{\pi}{2}} = \frac{1}{2}.$$

The amount of energy incident on unit-area in unit-time is hence equal to $\eta c/4$. Since this energy is completely absorbed by a black body, in the case of radiation equilibrium the emissive power must be equally large. Therefore we have

$$K = \frac{\eta c}{4}. \tag{7}$$

Since the modulus of distribution of the radiation is equal to $kT$, we have by § 155 eqn. (14)

$$\eta = aT^4, \tag{8}$$

and therefore

$$K = \frac{ac}{4}T^4, \tag{9}$$

where (because there are two degrees of freedom) the constant proportionality factor

$$a = \frac{48\pi\alpha k^4}{h^3c^3}. \tag{10}$$

The *emissive power of a black body* is therefore *proportional to the fourth power of the absolute temperature*. This important law was first derived in 1879 by *Stefan* from the experimental observations of other investigators, and was later proved by

*Boltzmann* on the basis of theoretical thermodynamical considerations (quite different, however, from those applied in the present derivation).[3] Measurements by *Kurlbaum* have given for the constant $a$ the value

$$a = 7{\cdot}06 \times 10^{-15} \frac{\text{erg.}}{\text{cm.}^3 \text{ deg.}^4}. \tag{11}$$

## § 175. Planck's Law of Radiation.

Like energy density, *emissive power* can also be *spectrally resolved* with respect to frequency intervals. The share taken in the total emissive power by the spectral region between $\nu$ and $(\nu + d\nu)$ is found from a combination of § 155 eqn. (6) with § 174 eqn. (7). Denoting this share by $K_\nu \, . \, d\nu$, we have, inasmuch as the number of degrees of freedom is two,

$$K_\nu = \frac{2\pi h \nu^3}{c^2} \cdot \frac{1}{e^{\frac{h\nu}{kT}} - 1}. \tag{1}$$

Now the spectrum as a whole can also be divided up according to wave-length, instead of to frequency. Let us consider the interval between a frequency $\nu'$ and a frequency $\nu''$, the corresponding wave-lengths being $\lambda'$ and $\lambda''$. Let $(\nu'' - \nu')$ be equal to $\Delta\nu$, and $(\lambda'' - \lambda')$ to $\Delta\lambda$. The share taken in the total emissive power by the small spectral region considered may then be equated to either $K_\nu \, \Delta\nu$ or $(-K_\lambda \, \Delta\lambda)$. (For the smaller wave-length corresponds to the larger frequency.) Now $\nu$ and $\lambda$ are connected by the elementary relation

$$\nu = \frac{c}{\lambda}. \tag{2}$$

Differentiating this, we find

$$d\nu = -\frac{c}{\lambda^2} d\lambda. \tag{3}$$

Hence

$$K_\lambda = \frac{c}{\lambda^2} K_\nu, \tag{4}$$

[3] Stefan's Law finds an interpretation from the point of view of the theory of light-quanta in the fact that the *number of light-quanta emitted* is *proportional to the cube of the temperature*, whilst the *mean energy* of an emitted light-quantum is proportional to the *temperature itself*. *Cf.* A. Haas, *Anzeiger Akad. Wiss. Wien, math.-nat. Klasse*, 1929, p. 65.

and therefore, by eqns. (1) and (2),

$$K_\lambda = \frac{2\pi h c^2}{\lambda^5} \cdot \frac{1}{e^{\frac{ch}{k\lambda T}} - 1}. \tag{5}$$

This equation is an expression of *Planck's Law of Radiation*. Planck derived it in 1900 by introducing into theoretical physics, as a radical innovation, the hypothesis of the *elementary quantum of action*. The confirmation of Planck's Radiation Law by accurate measurements constituted a brilliant justification for the new hypothesis, whereas all the earlier laws derived from the " classical theory " proved to be of only limited validity.[1]

Planck's Law determines the specific, or " monochromatic," emissive power of a small region of the spectrum as a function of two variables, the *temperature* and the *wave-length*. On differentiating partially with respect to the wave-length, we obtain the relation

$$\frac{\partial K_\lambda}{\partial \lambda} = -\frac{10\pi h c^2}{\lambda^6\left(e^{\frac{ch}{k\lambda T}} - 1\right)} + \frac{2\pi h c^2}{\lambda^5\left(e^{\frac{ch}{k\lambda T}} - 1\right)^2} \cdot \frac{ch}{k\lambda^2 T} \cdot e^{\frac{ch}{k\lambda T}}. \tag{6}$$

From this equation we can calculate the *wave-length* for which the *specific emissive power* represents a *maximum*. (That such a wave-length must exist follows from eqn. (5), for at every temperature $K_\lambda$ must vanish both when $\lambda = 0$ and when $\lambda = \infty$.) For this particular wave-length, which

[1] The two laws of radiation which had been previously derived form special cases of Planck's Law according as the product $\lambda T$ is taken to be small or large. In the first case, the exponential expression in eqn. (5) is large compared with unity. (It follows from the values of $h$ and $k$ that the expression $ch/k$ is approximately equal to $1\frac{1}{2}$.) Owing to the magnitude of the exponential expression, the number 1 may therefore be omitted in the special case considered, and eqn. (5) then assumes the special form

$$K_\lambda = \frac{2\pi h c^2}{\lambda^5} \cdot e^{-\frac{ch}{k\lambda T}}. \tag{5b}$$

This radiation law, which has thus only a *limited validity, viz.* in the *ultra-violet* or at *low temperatures*, was propounded by *Wien* in 1896.—If, on the other hand, the product $\lambda T$ is large, so that we are dealing with large wave-lengths or high temperatures, the value of the exponential in Planck's Radiation Law is small, and for these small values we can use the approximation formula in virtue of which, for small values of $x$, $e^x$ may be put equal to $(1 + x)$. Eqn. (5) then becomes

$$K_\lambda = \frac{2\pi c k T}{\lambda^4}. \tag{5c}$$

This equation is an expression of *Rayleigh's radiation law*, which is likewise of limited validity, holding *only* in the *infra-red* or at *high temperatures*. In passing it may be mentioned that it follows directly from Rayleigh and Jeans's principle, on the assumption that the mean value of the energy of a single characteristic vibration is equal to $kT$.

may be denoted by $\lambda_{max}$, $\partial K_\lambda/\partial\lambda$ must be equal to zero. If we put, for the sake of brevity,

$$\frac{c\,h}{k\,\lambda_{max}\,T} = \beta, \tag{7}$$

the relation

$$-5 + \frac{\beta e^{\beta}}{e^{\beta} - 1} = 0,$$

or

$$\frac{\beta}{5} + e^{-\beta} = 1 \tag{8}$$

must be satisfied. The quantity $\beta$ defined by eqn. (7) must thus be a constant, which is given as the root of the transcendental eqn. (8). Its value is

$$\beta = 4{\cdot}9651\,. \tag{9}$$

In accordance with eqn. (7), the product of the absolute temperature and the wave-length corresponding to the maximum of the specific emissive power must likewise be a *constant*. Denoting this constant by $b$, we may put

$$\lambda_{max}\,.\,T = b\,. \tag{10}$$

With increasing temperature the wave-length corresponding to the maximum specific emissive power is therefore displaced in the direction of smaller wave-lengths, and thus in the visible spectrum from the red to the violet end.[2] The law expressed by eqn. (10) was derived in 1893 on purely thermodynamic grounds by *Wien*, before the birth of the quantum theory, and it is known as *Wien's displacement law*. Exact measurements carried out by *Lummer* and *Pringsheim* have confirmed the law, and lead to a value for the constant $b$ of

$$b = 0{\cdot}294 \text{ cm. deg.} \tag{11}$$

According to eqn. (7), this constant is connected with the universal constants $h$ and $k$ by the simple relation

$$b = \frac{c\,h}{\beta\,k}. \tag{12}$$

[2] In this lies the explanation of the well-known experimental fact that, on heating, bodies first glow red; then at higher temperatures, yellow; and at very high temperatures, white. Moreover, the reason why all bodies begin to glow at the same temperature—525° C.—as *Draper* first observed, is explained by eqn. (10), for this contains no quantities depending on the chemical nature of the radiating body. It should, however, be expressly observed that Planck's and Wien's laws hold only for radiation which takes place at the cost of the heat content of the radiating body. (This is not the case with so-called luminescence phenomena.)

If this equation be combined with § 174 eqn. (10), a derivation of the quantities $h$ and $k$ ($= R/L$), and thus of the *elementary quantum of action* and of the *Loschmidt number*, from *radiation measurements* appears possible. This calculation was effected by *Planck* as long ago as 1900 (in the case of the quantum of action, with an inconsiderable error, but with an error of about 2 *per cent.* in the Loschmidt number).

This chapter may be brought to a close by summarizing in Table XXX. the most important of the manifold methods

TABLE XXX

*Methods of Derivation of the Fundamental Constants of Atomistics*

| Method. | Reference. | Relevant combination between $e$, $h$, $m$, and $L$. |
|---|---|---|
| 1. Individual observation of small electric charges | § 86 eqn. (7) | $e$ |
| 2. Electrolysis | § 87 eqn. (5) | $L\,e$ |
| 3. Deflection of cathode rays | § 88 eqn. (4) | $e/m$ |
| 4. Deflection of α-rays | § 89 eqn. (1) | $L\,e$ |
| 5. Scintillation of α-particles | § 89 | $e$ |
| 6. Photo-electric effect | § 90 eqn. (2) | $e/h$ |
| 7. Rydberg's constant | § 104 eqn. (18) | $e^4 m / h^3$ |
| 8. Fine structure of spectral lines | § 104 eqns. (19) and (26) | $e^2/h$ |
| 9. Displacement towards the violet of helium lines | § 105 eqn. (26) | $L\,m$ |
| 10. Zeeman effect | § 113 eqn. (1) | $e/m$ |
| 11. Stark effect | § 114 eqn. (2) | $h/e\,m$ |
| 12. Formation of helium from radium | § 124 | $L$ |
| 13. Excitation of X-rays | § 142 | $e^{4/3}/h$ |
| 14. Brownian movement | § 169 | $L$ |
| 15. Sedimentation equilibrium | § 170 | $L$ |
| 16. Colour of sky | § 170 | $L$ |
| 17. Stefan's Law | § 174 eqn. (10) | $h^3\,L^4$ |
| 18. Wien's Displacement Law | § 175 eqn. (12) | $L\,h$ |

which permit of a *calculation of the four fundamental constants of atomistics, viz.*, the elementary quantum of action, the Loschmidt number, the elementary quantum of electricity, and the mass of the electron.[3]

In whatever way the four unknowns are derived according to these various methods, the same values (within the limits of error) are always obtained. If the fact that the sensations of sight, hearing, and taste all indicate the same objects is to be taken as an argument for the existence of the world around us, then theoretical physics may regard as a *certain proof of the actual existence of electrons, atoms, and the elementary quantum of action* the fact that the *same values* are always obtained for the characteristic constants of atomistics, even by fundamentally different methods.

[3] In the case of method 13 in the table, it should be remembered that the value of the Loschmidt number is implicitly contained in the wave-lengths of the X-rays [according to § 120 eqn. (3) ], and therefore, when Faraday's constant ($L\,e$) is known, the value of the elementary quantum of electricity is likewise so contained.

# PART V

## THEORY OF RELATIVITY

## Chapter XXI

# SPECIAL THEORY OF RELATIVITY

### § 176. Michelson's Experiment.

In § 13 it was shown that the theory of relative motion leads to the *mechanical principle of relativity*, according to which all mechanical equations have the same form for two coordinate systems *moving uniformly* with respect to each other. Moreover, it must appear inherently improbable that the same principle should not also apply to the equations of *electrodynamics* and *optics*, but a brief consideration will show that such a supposition meets with great difficulties.

Let us consider two parallel coordinate systems, of which the second is moving with a velocity **u** with respect to the first (in a direction coinciding with the common $x$-direction). Distinguishing the coordinates in the first system by single dashes, and those in the second system by double dashes, then in accordance with the usual laws of kinematics the coordinates in the two systems will be connected by the relations

$$x'' = x' - u\,t,\quad y'' = y',\quad z'' = z'\,. \tag{1}$$

Hence, if the form of a mechanical equation in the first system be

$$\varphi\,(x', y', z', t) = 0\,, \tag{2a}$$

it follows from the mechanical principle of relativity that the *same* function must also be zero when $x'$ is replaced by $(x' - u\,t)$, so that

$$\varphi\,(x'', y'', z'', t) = 0\,. \tag{2b}$$

We see at once, however, that the equations, the form of which remains unaltered on the transition from the first to the second coordinate system, *cannot include* the equation describing the *uniform propagation of light in spherical waves*. If this equation, which states that

$$x^2 + y^2 + z^2 - c^2\,t^2 = 0\,, \tag{3}$$

holds good in the first coordinate system, it cannot do so in

the second, for, when $u$ differs from zero, it follows inevitably from eqn. (1) that

$$(4) \quad x''^2 + y''^2 + z''^2 = c^2 t^2 \neq x'^2 + y'^2 + z'^2 - c^2 t^2 .$$

In contradistinction to the mechanical principle of relativity, which puts the idea of an absolute motion beyond the bounds of possibility, the above inequality appears to demand the existence of *only one single system* in which light is propagated uniformly in all directions with the same velocity. If this were the case, however, there could evidently be no question of a uniform propagation of light in all directions on the moving earth. A confirmation of inequalities in this propagation thus presented to physicists the enticing possibility of demonstrating, or even of actually measuring, the motion of the earth relatively to that privileged single system, *i.e.*, the *motion of the earth relatively to the so-called æther*.

Among the manifold experiments undertaken with this object [1] in view, by far the greatest importance attaches to the famous *experiment of Michelson* (1881). In this experiment, a ray of light from a source $L$ (Fig. 132) strikes a half-silvered mirror $S$ at an angle of 45°. In this way it is split into two mutually perpendicular rays which are reflected by two other mirrors, $S_1$ and $S_2$, the distances between these and $S$ being approximately the same ($l$). At $P$ the interference phenomena due to the superposition of the rays coming from the mirrors $S_1$ and $S_2$ are observed through a telescope. For, since the intervals $SS_1$ and $SS_2$ differ, in general, by a very small amount, the same equidistant

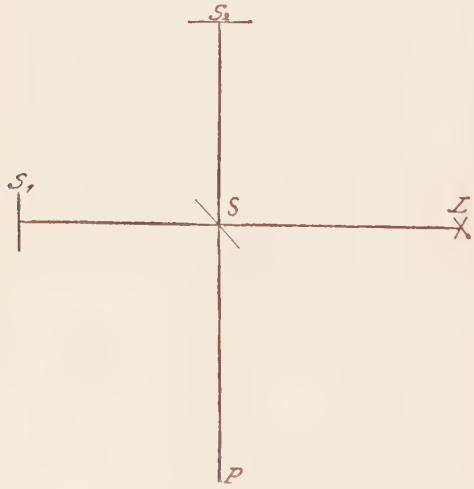


FIG. 132.

[1] An extensive summary of all these experiments is to be found in the first chapter of "*Relativitätstheorie*," by M. v. Laue (3rd edition; Brunswick, 1921).

interference fringes will be observed in the telescope as are produced, with suitable apparatus, by reflection at plane-parallel plates (we cannot enter into further details here).

On the assumption that an æther really exists, in which light is propagated in all directions with the same velocity $c$, we shall suppose the translatory velocity of the earth relatively to this æther to be $\mathbf{u}$. Let $\mathbf{v}_a$ be the velocity of a light ray of given direction with respect to the æther, and $\mathbf{v}_e$ its velocity with respect to an observer on the earth. Then, according to the formula for relative velocity (and neglecting the rotation of the earth),

$$\mathbf{v}_a = \mathbf{v}_e + \mathbf{u}\,, \tag{5}$$

where the magnitude of the vector $\mathbf{v}_a$ must always be the same ($c$).

We shall now suppose the apparatus to be so arranged that the straight line $SS_1$ lies in the direction of the vector $\mathbf{u}$. We shall denote the magnitude of the velocity of light (as observed on the earth) along the interval $SS_1$ by $v_1'$, and along the interval $S_1S$ by $v_1''$. Since, in these two special cases, the vectors $\mathbf{v}_e$ and $\mathbf{u}$ are similarly or oppositely directed respectively, we obtain from eqn. (5) the relations

$$\begin{cases} c = v_1' + u\,, \\ c = v_1'' - u\,. \end{cases} \tag{6}$$

Let $v_2'$ be the velocity of light with respect to the earth along the interval $SS_2$, and $v_2''$ the velocity along the interval $S_2S$. Since the vectors $\mathbf{v}_e$ and $\mathbf{u}$ are mutually perpendicular in these two special cases, it follows from the Theorem of Pythagoras that the square of the magnitude of their geometrical sum is simply equal to the sum of the squares of the individual magnitudes. Hence we find from eqn. (5) the relations

$$\begin{cases} c^2 = v_2'^2 + u^2\,, \\ c^2 = v_2''^2 + u^2\,. \end{cases} \tag{7}$$

If we denote by $t_1$ the time required by the light to travel along the path $SS_1$ and back, and by $t_2$ the time required to travel along the path $SS_2$ and back, we have, by eqns. (6) and (7),

$$t_1 = \frac{l}{v_1'} + \frac{l}{v_1''} = \frac{l}{c-u} + \frac{l}{c+u} = \frac{2l}{c}\left(1 - \frac{u^2}{c^2}\right)^{-1}, \tag{8}$$

and

$$t_2 = \frac{l}{v_2'} + \frac{l}{v_2''} = \frac{2l}{\sqrt{c^2 - u^2}} = \frac{2l}{c}\left(1 - \frac{u^2}{c^2}\right)^{-\frac{1}{2}}. \tag{9}$$

With the aid of the binomial theorem, we can develop eqns. (8) and (9) into series. For that purpose we use the formula

$$(10) \qquad (1+x)^n = 1 + \frac{n}{1}x + \frac{n(n-1)}{1\cdot 2}x^2 + \ldots\ldots,$$

which is generally valid provided that

$$-1 < x < +1.$$

According to this formula,

$$(11) \qquad \left(1-\frac{u^2}{c^2}\right)^{-1} = 1 + \frac{u^2}{c^2} + \frac{u^4}{c^4} + \ldots\ldots,$$

and

$$(12) \qquad \left(1-\frac{u^2}{c^2}\right)^{-\frac{1}{2}} = 1 + \frac{1}{2}\frac{u^2}{c^2} + \frac{3}{8}\frac{u^4}{c^4} + \ldots\ldots$$

Hence, as a first approximation,

$$(13) \qquad t_1 - t_2 = \frac{l}{c}\frac{u^2}{c^2}.$$

Now if, in consequence of the earth's motion, the quantity $(t_1 - t_2)$ differed from zero, this would obviously result in a *displacement of the interference fringes* relatively to the position which they would occupy, were the earth at rest with respect to the æther. Let $r'$ be the magnitude of this displacement, $r$ the distance between two adjacent interference fringes, and $\tau$ the period of the monochromatic light used; then the relation

$$(14) \qquad r' : r = (t_1 - t_2) : \tau$$

must evidently exist. Hence, denoting the wave-length of the light by $\lambda$, and remembering that $\lambda$ is equal to $c\,\tau$, we have

$$(15) \qquad \frac{r'}{r} = \frac{l}{\lambda}\cdot\frac{u^2}{c^2}.$$

If now we *rotate* the whole apparatus *through a right angle* in its own plane, the time $t_1$ required by the light to travel along the path $SS_1$ and back again then becomes equal to the previous value $t_2$. On the other hand, the time $t_2$ required for traversing the interval $SS_2$ and back again is now equal to the original $t_1$. Hence we obtain for the difference $(t_1 - t_2)$ the amount

$$(16) \qquad t_1 - t_2 = -\frac{l}{c}\frac{u^2}{c^2}.$$

The displacement of the interference fringes with respect to the positions which they would occupy were the earth at rest must now be equal, but opposite in sense, to the original displacement. In consequence of the rotation of the apparatus, the interference fringes must therefore appear displaced by an interval $2r'$ relatively to their position with the original orientation of the apparatus. If we express this total displacement due to the rotation of the apparatus as a fraction ($\alpha$) of the interval between two adjacent fringes, we have therefore

$$\alpha = 2\frac{l}{\lambda}\frac{u^2}{c^2}. \tag{17}$$

Admittedly, the translatory velocity of the earth is only the ten-thousandth part of the velocity of light, so that $u^2/c^2$ is equal to $10^{-8}$. On the other hand, however, the length $l$ in Michelson's experiment measured a few metres, and, in a later and improved experiment carried out by Michelson and *Morley* (1887), was increased to 11 m.[2] Since $\lambda$ is approximately $6 \times 10^{-5}$ cm., the resulting value of $\alpha$ was 0·4. The rotation of the apparatus should thus have been accompanied by a displacement of the interference fringes through about half the distance between two neighbouring fringes. Of the expected and quite considerable displacement, however, nothing was to be observed.[3]

## § 177. Einstein's Principle of Relativity.

Michelson and Morley's experiment necessarily led to the conclusion that, despite the motion of the earth, light is nevertheless propagated with the same velocity in all directions relatively to a terrestrial observer. Thus the *principle of relativity* does actually appear to be as valid for *electrodynamical and optical phenomena* as it is for mechanical phenomena. *All equations* which describe any kind of physical process with reference to a definite system must, therefore, *retain the same form* when the original system is replaced by another *moving uniformly* with respect to it. We are led to this conclusion not only by Michelson's experiment; for it complies with the concept of physical

[2] Each ray of light was reflected several times to and fro.

[3] The great importance of Michelson's experiment lies in the fact that the expected effect depends on the *square* of the ratio $u/c$. Lorentz, indeed, had previously shown (we cannot enter into further details here) that the absence of any effect could always be explained on the basis of simple assumptions, if it depended only on the first power of the ratio $u/c$.

uniformity, with which any limitation of the validity of the relativity principle to a single branch of physics would be incompatible. On the other hand, however, an extension of that principle to the fields of electrodynamics and optics must necessarily stand in contradiction to the inequality expressed by § 176 (4); and thus theoretical physics was faced with the difficult problem of solving this incompatibility.

The first attempt at a solution was made by *Lorentz* and *Fitzgerald*, who assumed that a body moving through the æther experiences a *contraction* in the direction of its motion and depending on the magnitude of its velocity, and this *ad hoc* hypothesis certainly afforded an explanation of the negative result of Michelson's experiment. It was, however, due to the critical ingenuity of *Einstein* that the contradiction was first solved from a much more general point of view. Einstein accomplished this by liberating physics from a deeply rooted prejudice which, owing principally to *Newton's* influence, had become a fundamental concept of theoretical physics; this was the concept of *absolute time*.[1]

Einstein pointed out that it was entirely arbitrary and unjustifiable to ascribe the same scales of time to two systems moving relatively to each other. On the contrary, it must be postulated that observers in the two systems shall use *different time-scales* for their measurements of time, exactly as they would use different spatial coordinates in stating the position of any given body.

Thus, whereas the two sides of the inequality (4) in § 176 must always differ so long as the concept of an absolute time is retained, a transformation of the inequality into an equation becomes possible when on the left-hand side $t'$, and on the right-hand side $t''$, is substituted for $t$; $t'$ and $t''$ are the times obtained in measurements by the observers in the first and second systems respectively. The simple relation between $x''$ and $x'$ expressed by § 176 eqn. (1) must then, of course, be replaced by a more complicated relation, in which absolute time will not appear, as it did in the original equation.

By uprooting, in this way, the prejudice of an absolute time and demolishing the great obstacle to an extension of the concept of relativity to the whole territory of physics,

[1] The first definition in Newton's "*Philosophiæ naturalis principia mathematica*" runs: "Tempus absolutum verum et mathematicum in se et natura sua sine relatione ad externum quodvis æqualiter fluit." Newton's conception of absolute motion has been resolutely opposed, notably by *Mach*.

Einstein was able, in 1905, to propound his famous so-called *Special Principle of Relativity*. According to this principle, every physical event which is described by a definite equation in a definite reference system must be described, in any other system *moving uniformly and rectilinearly* with respect to the first, by the *same equation*, provided that the spatial and temporal coordinates of the first system be replaced by those of the second.

Conversely, the spatial and temporal coordinates in two systems moving uniformly and rectilinearly with respect to each other must be connected, quite generally, by a *universal relation depending only on the mutual relative velocity*. The form of this relation is determined by the postulate that, *in both systems*, the *velocity of light* shall appear to have the *same value in all directions*, this value being a universal constant.

## § 178. The Minkowski-World.

Einstein's conception of the relativity of time renders possible a transformation into an equation of the inequality expressing the apparent incompatibility of Michelson's experiment with classical theoretical physics [§ 176 formula (4)]. Hence it permits also of an extension of the validity of the relativity principle from mechanics to the fields of electrodynamics and optics. Absolute time ($t$) is replaced by the relative time ($t'$) for the first system, and by the relative time ($t''$) for the second system. The sign of inequality becomes the symbol for an equation; and the inequality itself is replaced by the *fundamental relation of the theory of relativity*:

$$(1)\quad x''^2 + y''^2 + z''^2 - c^2 t''^2 = x'^2 + y'^2 + z'^2 - c^2 t'^2 .$$

In 1908 this equation was given a *geometrical interpretation* by *Minkowski*,[1] and this allows of a considerable simplification of all deductions made from Einstein's Principle of Relativity. Let us consider for a moment the analytical geometry of a plane surface; it is then a property of the *sum of the squares of the coordinates* $x$ and $y$ that its *value* is the *same* for all arbitrary rectangular coordinate systems based on the same origin. This sum is found to be independent of the orientation of the coordinate system. Analogous considerations apply to the sum of the squares

[1] A paper published by Poincaré in 1906 touched, indeed, on Minkowski's concepts.

of the *three* coordinates $x$, $y$, $z$ in a three-dimensional space. Hence, as Minkowski recognized, Einstein's Principle of Relativity as expressed by eqn. (1) may also be interpreted geometrically as follows: The square of the product of the *time* and the velocity of light, taken with a negative sign, and the squares of the three *spatial* coordinates stand in the same relation to each other as do the squares of the *four coordinates in a four-dimensional geometry.*

The four-dimensional manifold, which appears as a *linking together of space and time* by these four coordinates, is known as the *Minkowski-world.* The " world-coordinates " are therefore the four quantities $x$, $y$, $z$, and $i\,c\,t$ (where $i$, as usual, represents $\sqrt{-1}$). The location of a moving particle at a given time $t^*$ at a given point $(x^*, y^*, z^*)$ may thus be expressed by the fixation of a so-called *world-point*, the coordinates of which are equal to $x^*$, $y^*$, $z^*$, and $i\,c\,t^*$. Quantities which can be represented by directed intervals in the Minkowski-world are called *world-vectors.*

The *motion* of a particle may thus be represented by a so-called *world-line*, which expresses the dependency of the path on the time, whilst its projection in three-dimensional space gives the course of the path.[2] The form of a world-line and the length of its elements are therefore independent of the reference system. If we denote the length of a world-line-element by $ds$, we have (by an extension of Pythagoras' Theorem) for an arbitrary reference system the equation

$$ds^2 = dx^2 + dy^2 + dz^2 - c^2\,dt^2\,. \tag{2}$$

Again, the *velocity* ($v$) of the particle, the motion of which is represented by the world-line, with respect to the system used is given by the equation

$$v^2 = \left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2. \tag{3}$$

Hence

$$ds^2 = (v^2 - c^2)\,dt^2\,, \tag{4}$$

and therefore, since $1/i = -i$,

$$\frac{dt}{ds} = -\frac{i}{c\sqrt{1 - \dfrac{v^2}{c^2}}}\,. \tag{5}$$

[2] This is in no way to be taken as implying that, conversely, every world-line may be regarded as the representation of a motion. *Cf.* § 181, footnote 1.

Finally, let us see what is the meaning of the four *directional cosines* associated with a *world-line-element*; we shall denote them by $\alpha_1, \alpha_2, \alpha_3, \alpha_4$. We have

$$\alpha_1 = \frac{dx}{ds} = \frac{dx}{dt} \cdot \frac{dt}{ds},$$

where $dx/dt$ is equal to $v_x$. On the other hand, inasmuch as the fourth coordinate is $i\,c\,t$,

$$\alpha_4 = i\,c\,\frac{dt}{ds}.$$

On substituting for $dt/ds$ its value as given by eqn. (5), we obtain

$$(6) \qquad \alpha_1 = -\frac{i}{c}\frac{v_x}{\sqrt{1-\frac{v^2}{c^2}}}, \qquad \alpha_4 = \frac{1}{\sqrt{1-\frac{v^2}{c^2}}}.$$

The formulæ for $\alpha_2$ and $\alpha_3$ are similar to that for $\alpha_1$, except that $v_x$ is replaced by $v_y$ or $v_z$ respectively. Conversely, we can derive from eqns. (6) the velocity of the particle with respect to the system used, when the cosines of the angles included between the world-line-element of the particle and the coordinate axes are given as equal to $\alpha_1, \alpha_2, \alpha_3, \alpha_4$. The *magnitude* of the velocity follows directly from the value of $\alpha_4$ alone.

If, in particular, the four-dimensional coordinate system be so arranged that its *time-axis coincides* with the direction of the *world-line*, we have in respect to this system

$$(7) \qquad \alpha_1 = \alpha_2 = \alpha_3 = 0, \quad \alpha_4 = 1.$$

In accordance with eqns. (6) $v$ is then equal to 0. Relatively to this system, the particle is therefore *at rest*.

## § 179. The Lorentz-Transformation.

When we pass from one four-dimensional coordinate system $(x, y, z, i\,c\,t)$ to a second $(x', y', z', i\,c\,t')$, four equations govern the accompanying *transformation* [in accordance with § 9 eqns. (1) and (3)]. We shall, however, write down only two of these, *viz.*,

$$(1) \qquad \begin{cases} x' = x\cos(x', x) + y\cos(x', y) + z\cos(x', z) \\ \qquad\qquad + i\,c\,t\cos(x', t), \\ i\,c\,t' = x\cos(t', x) + y\cos(t', y) + z\cos(t', z) \\ \qquad\qquad + i\,c\,t\cos(t', t). \end{cases}$$

We shall only consider the simple special case in which

the $y$- and $z$-axes are common to both four-dimensional coordinate systems; the transformation from one system to the other is then produced by a *rotation in the $x$-$t$-plane*. If $\varphi$ be the angle through which the $x$-axis must be turned in order to make it coincide with the $x'$-axis, the angle between $t$ and $t'$ is also equal to $\varphi$. The angle between $x'$ and $t$ will be equal to $(90^\circ - \varphi)$, so that its cosine is equal to $\sin \varphi$, and the angle between $x$ and $t'$ will be $(90^\circ + \varphi)$, so that its cosine is equal to $(-\sin \varphi)$. Of the other twelve cosines occurring in the four eqns. (1) $\cos(y', y)$ and $\cos(z', z)$ are equal to unity, whilst all the others are zero. In the special case considered, therefore, eqns. (1) (on division of the last equation by $i\,c$) assume the simpler form

$$(2) \qquad \begin{cases} x' = x\cos\varphi + i\,c\,t\sin\varphi\,, \\ y' = y, \quad z' = z\,, \\ t' = \dfrac{i}{c}\,x\sin\varphi + t\cos\varphi\,. \end{cases}$$

If, now, we consider a particle which is at rest with respect to the coordinate system $(x', y', z')$, it follows from the concluding remarks of § 178 that its world-line must coincide with the $t'$-axis. With respect to the coordinate system $(x, y, z)$, however, it will be moving with a velocity $\mathbf{v}$, which can be determined from § 178 eqns. (6) when we put

$$(3) \qquad \begin{cases} \alpha_1 = \cos(t', x), \quad \alpha_2 = \cos(t', y)\,, \\ \alpha_3 = \cos(t', z), \quad \alpha_4 = \cos(t', t)\,. \end{cases}$$

Since the angles between $t'$ on the one hand and $y$ and $z$ on the other are right angles, $v_y$ and $v_z$ vanish, so that $v_x$ is equal to $v$. Inasmuch as $\alpha_1$ is equal to $(-\sin\varphi)$, and $\alpha_4$ to $\cos\varphi$, we have $\alpha_1/\alpha_4$ equal to $(-\tan\varphi)$. Hence, introducing the values given in § 178 eqns. (6) for $\alpha_1$ and $\alpha_4$, we find

$$(4) \qquad \tan\varphi = i\frac{v}{c}\,.$$

Again, it follows directly from § 178 eqns. (6) [seeing that $\alpha_1$ is equal to $(-\sin\varphi)$, and $\alpha_4$ to $\cos\varphi$], or from eqn. (4) with the help of elementary trigonometrical formulæ, that

$$(5) \qquad \begin{cases} \sin\varphi = \dfrac{i}{c}\dfrac{v}{\sqrt{1 - \dfrac{v^2}{c^2}}}\,, \\ \cos\varphi = \dfrac{1}{\sqrt{1 - \dfrac{v^2}{c^2}}}\,. \end{cases}$$

Now if, as we have assumed, a particle is at rest with respect to the second coordinate system, whilst with respect to the first it is moving with the velocity $v$ in the common $x$-direction, this means that the *spatial coordinate system* $(x', y', z')$ is *moving with a velocity* $v$ in the direction of the common $x$-axis relatively to the *spatial coordinate system* $(x, y, z)$. A *rotation of four-dimensional coordinate axes* through an angle $\varphi$ in the $x$-$t$-plane is therefore *equivalent to a transformation from one spatial coordinate system to another* which is *moving uniformly* in the common $x$-direction *with a velocity $v$ relatively to the first system*. In this rotation, the tangent of the angle $\varphi$ is equal [according to eqn. (4)] to the quotient of the velocity $v$ and the velocity of light, multiplied by the imaginary quantity $i$.

Eqns. (2) and (5) enable us at once to answer the question as to how the *spatial coordinates and the time are transformed* in accordance with the *Principle of Relativity*, when we pass from a given spatial coordinate system to another system parallel to the first and *moving uniformly* relatively to it, with a velocity $v$ in a direction which we suppose to coincide with the common $x$-axis. For, substituting in eqns. (2) the values given by eqns. (5), we find

$$(6) \qquad \begin{cases} x' = \dfrac{x - v\,t}{\sqrt{1 - \dfrac{v^2}{c^2}}}, \\ y' = y, \quad z' = z, \\ t' = \dfrac{-\dfrac{v}{c^2}\,x + t}{\sqrt{1 - \dfrac{v^2}{c^2}}}. \end{cases}$$

The totality of the substitutions expressed by eqns. (6) are known as a *Lorentz-transformation* [because the first attempts at deriving eqns. (6) were made by Lorentz in a paper published in 1904]. If, in particular, the *velocity $v$ is small as compared with the velocity of light*, the Lorentz-transformation assumes the special form

$$(7) \qquad x' = x - v\,t, \quad y' = y, \quad z' = z, \quad t' = t.$$

In three of these relations we recognize the substitutions which, according to the previously quoted laws of so-called "classical" kinematics, connect the coordinates of two systems moving uniformly with respect to each other. These laws were stated in § 176 eqns. (1), and the fourth relation

was taken for granted in classical mechanics, the same time-scale being assigned to both systems.[1]

Since of two spatial coordinate systems moving uniformly with respect to each other neither, of course, is in any way to be preferred to the other, it must be possible, quite generally, to derive the "undashed" variables $(x, y, z, t)$ from the "dashed" by means of a Lorentz-transformation, thus by means of eqns. (6), provided only that the quantity $(+ v)$ be replaced by the quantity $(- v)$. For, if the second system be moving with a velocity $v$ relatively to the first, the relative translational velocity of the first with respect to the second system must be $(- v)$ in the common $x$-direction.

On forming the sum of the squares of the four "dashed" world-coordinates, we find by eqns. (6)

$$x'^2 + y'^2 + z'^2 - c^2 t'^2 = y^2 + z^2 + \frac{x^2 - 2xvt + v^2 t^2 - \frac{v^2 x^2}{c^2} + 2xvt - c^2 t^2}{1 - \frac{v^2}{c^2}}$$

$$= y^2 + z^2 + \frac{x^2\left(1 - \frac{v^2}{c^2}\right) - t^2(c^2 - v^2)}{1 - \frac{v^2}{c^2}} = x^2 + y^2 + z^2 - c^2 t^2 .$$

Bearing in mind § 178 eqn. (1), we see that the Lorentz-transformation does actually satisfy Einstein's Principle of Relativity.

In point of fact, the equations expressing the Lorentz-transformation can be directly derived from § 178 eqn. (1), without having to introduce the Minkowski-world. The calculations, however, are then much more complicated and tedious than those just performed.

## § 180. Relativity of Scales of Time and Length.

Let us consider, with reference to the undashed system, two events, one of which occurs at the point $(x_1, y_1, z_1)$ at the time $t_1$, and the other at the point $(x_2, y_2, z_2)$ at the time $t_2$. Then, in accordance with the equations expressing the Lorentz-transformation, the times at which the two

[1] The special form of the Lorentz-transformation expressed by eqn. (7) is known as a Galileo-transformation, because it leads to the formulæ of classical mechanics established by Galileo.

events occur relatively to the dashed system are given by the formulæ

$$(1) \qquad t_1' = \beta\left(t_1 - x_1\frac{v}{c^2}\right), \quad t_2' = \beta\left(t_2 - x_2\frac{v}{c^2}\right),$$

where, for brevity,

$$(2) \qquad \beta = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

If, in particular, we assume that

$$t_1 = t_2,$$

it does not in any way follow that the quantities $t_1'$ and $t_2'$ are identical, provided that $x_1$ and $x_2$ differ. For an observer in the undashed system the two events, which occur at different places, are *simultaneous;* but to an observer in the dashed system they appear to happen *at different times.* In general, therefore, two events which are simultaneous in reference to one system are not simultaneous when referred to another system moving relatively to the first. Thus, in the theory of relativity, the concept of *simultaneity*, like that of time, has *not an absolute, but only a relative meaning.*

When, in particular, we put

$$x_1 = x_2$$

in eqns. (1), this means that both events occur at the same place for an observer in the undashed system. In this special case, it follows from eqns. (1) and (2) that

$$(3) \qquad t_2' - t_1' = \frac{t_2 - t_1}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

Since the denominator of the fraction is always smaller than unity, the intervals of time between the two events appears larger to the observer in the dashed system than to the observer in the undashed system. Conversely, among all systems in a state of uniform relative motion, that one in which the two events occur at the same place is characterized by the fact that the interval between the two events appears shortest to an observer in that system. A process which, with reference to a given system, occurs at one place alone (which, as we may say, is "fixed in space" with respect to this system) appears to run its course *most rapidly* to an observer in this system, but *more slowly* to an observer in any other. The process, indeed, will appear the slower the

more quickly that any system is moving relatively to the first system.[1] This conclusion from the Principle of Relativity is commonly stated in the form of a concise, but not very clearly expressed, rule, *viz.*, that a moving clock goes more slowly than one at rest.

The difference in time-scales for two systems in relative motion suggests that we should *compare scales of length* in the two systems. Now the *length of a rod* which is moving in an arbitrary system can, in general, evidently be determined only by measuring the coordinates $(x_1, y_1, z_1)$ of one end of the rod at an arbitrary time $t_1$, and *simultaneously* (or, more accurately, at a time $t_2$ which, with respect to the system, is identical with $t_1$) measuring the coordinates $(x_2, y_2, z_2)$ of the other end. The length of the rod in the system is then determined by the relation

$$l = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}\,. \tag{4}$$

If, as we shall assume, the rod always remains parallel to the $x$-axis, the equation becomes simply

$$l = x_2 - x_1\,. \tag{5}$$

We shall now so arrange the undashed coordinate system that, with respect to it, the rod is *at rest* and lying in the direction of the $x$-axis. Then, since $x_1$ and $x_2$ are invariable, it is obviously quite unnecessary that the times $t_1$ and $t_2$, at which the coordinates $x_1$ and $x_2$ are measured, should be the same. These times may be quite arbitrary without the value obtained for $l$ being in any way affected.

If, next, we pass from the undashed to the dashed system, which is connected with the former by a Lorentz-transformation, then with respect to the latter system, relatively to which the rod is moving with a velocity of magnitude $v$,[2] the length of the rod is evidently given by the relation

$$l' = x_2' - x_1'\,. \tag{6}$$

In this equation $x_1'$ and $x_2'$ are the dashed coordinates corresponding to the undashed coordinates $x_1$ and $x_2$ in accordance with the Lorentz-transformation, whilst $t_1'$ and $t_2'$ must be identical. Now

$$x_1' = \beta\,(x_1 - v\,t_1), \quad x_2' = \beta\,(x_2 - v\,t_2), \tag{7}$$

[1] It must be remembered that, in accordance with eqn. (3), the degree of retardation is independent of the sign of $v$.

[2] The vectorial value of this relative velocity is $(-\mathfrak{v})$.

so that

$$x_2' - x_1' = \beta\,(x_2 - x_1) - \beta\,v\,(t_2 - t_1)\,. \tag{8}$$

Again, since $t_2'$ and $t_1'$ must be identical, it follows from the last of § 179 eqns. (6) that

$$-\frac{v}{c^2}\,x_1 + t_1 = -\frac{v}{c^2}\,x_2 + t_2\,,$$

or,

$$t_2 - t_1 = \frac{v}{c^2}\,(x_2 - x_1)\,. \tag{9}$$

Introducing this value into eqn. (8), we obtain

$$x_2' - x_1' = \beta\,(x_2 - x_1)\left(1 - \frac{v^2}{c^2}\right),$$

or, by eqn. (2),

$$l' = l\,\sqrt{1 - \frac{v^2}{c^2}}\,. \tag{10}$$

Therefore to an observer who sees the rod *in motion*, it appears *shortened*;[3] and, among all systems in a state of uniform relative motion, that one in which the rod is at rest is distinguished from all others by the fact that, in it, the rod appears longer than in any other system.[4] For example, a metre rod lying on the earth in the direction of its motion would appear to an observer on the sun to be shortened by $5 \times 10^{-7}$ cm.

If the direction of the rod be that of the $y$- or the $z$-axis, its length will be the same with respect to both the dashed and the undashed system, for $y' = y$ and $z' = z$, and the time-coordinates are independent of $y$ and $z$, or of $y'$ and $z'$, as the case may be. A body which is at rest and appears spherical to an observer in the undashed system must hence appear as an *oblate spheroid* to an observer in the dashed system. The theory of relativity recognizes *no absolute*, but only a *relative shape*.

## § 181. Composition of Velocities.

Minkowski's representation enables us to supply a ready answer to the question as to the *composition of velocities* from the point of view of the relativity theory; in the following

[3] The *contraction hypothesis* of Lorentz and Fitzgerald (*cf.* § 177) thus actually appears as a necessary consequence of Einstein's Principle of Relativity.

[4] As eqn. (10) shows, the contraction is independent of the sign of $v$. A rod which is at rest in the dashed system will appear shortened to an observer in the undashed system, in exactly the same way as a rod at rest in the undashed system does to an observer in the dashed system.

discussion, however, we shall only consider the especially simple case in which the two component velocities have the same direction. Let $v'$ be a velocity with respect to a system which is itself moving in the same direction with a velocity $v$ relatively to a standard system ; we wish to find the resultant velocity ($u$) with respect to the latter. We shall suppose the common direction of the velocities $v$, $v'$, and $u$ to coincide with the common $x$-axis of the coordinate systems.

We shall denote the original coordinate system (in which the velocity is $u$) by I ; that in which the velocity is $v'$ by II ; and, finally, one in which the velocity vanishes by III. Now if the velocity vanishes in III, whilst in I it is equal to $u$, it follows from § 179 eqn. (4) that the relation

$$i\frac{u}{c} = \tan\psi \tag{1}$$

must exist, where $\psi$ is the angle through which system I must be rotated in the $x$-$t$-plane in order to obtain system III.

On the other hand, however, we can also pass from system I to system III by first transforming to system II, which is moving with a velocity $v$ with respect to system I, and thence to system III, which is moving with a velocity $v'$ relatively to system II. In a four-dimensional representation the first transformation appears as a rotation through an angle $\varphi$, and the second as a rotation through an angle $\varphi'$, where

$$i\frac{v}{c} = \tan\varphi\,, \qquad i\frac{v'}{c} = \tan\varphi'\,. \tag{2}$$

The transformation from I to III is therefore represented as a sequence of two rotations through the angles $\varphi$ and $\varphi'$, and hence

$$\psi = \varphi + \varphi'. \tag{3}$$

But, according to an elementary trigonometrical formula,

$$\tan(\varphi + \varphi') = \frac{\tan\varphi + \tan\varphi'}{1 - \tan\varphi\tan\varphi'}\,.$$

Hence, by eqns. (1), (2), and (3),

$$i\frac{u}{c} = \frac{i\frac{v}{c} + i\frac{v'}{c}}{1 - i^2\frac{vv'}{c^2}}\,,$$

or,

$$u = \frac{v + v'}{1 + \frac{v\,v'}{c^2}}. \tag{4}$$

For the special case in which $v'$ is equal to $c$, we have

$$u = \frac{v + c}{1 + \frac{v}{c}} = c. \tag{5}$$

Thus, according to the theory of relativity, the addition of an arbitrary velocity to the velocity of light always gives the quantity $c$ again. Hence the *velocity of light* plays a similar part in the theory of relativity to that played by *infinity* in mathematics. It appears from the relativity theory that *velocities larger than that of light* are *impossible*.[1]

[1] From the impossibility of velocities exceeding that of light it follows also that a *world-vector* drawn between two world-points can *only represent a motion* when the *square of the distance* between the two world-points is *negative*. For, if the world-vector is to represent a motion, the relation

$$\Delta r < c\,\Delta t \tag{6}$$

must hold, where $\Delta r$ denotes the spatial interval between the two world-points as defined by the equation

$$(\Delta r)^2 = (x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2. \tag{7}$$

Furthermore,

$$(\Delta s)^2 = (\Delta r)^2 - c^2(\Delta t)^2. \tag{8}$$

Using a coordinate system in which the particle is at rest, we can take the world-vector as our *time-axis*, so that $\Delta r$ then vanishes, and the two world-points joined by the world-vector *coincide in their positions in space*. If we pass from the reference system in which the time-axis coincides with the world-vector to any other system moving with a velocity $v'$ relatively to the former, we have in this latter system

$$(\Delta r'^2) = v'^2(\Delta t')^2. \tag{9}$$

Hence, since $v'$ cannot be larger than $c$, $(\Delta s)^2$ must be *negative in every reference system* if the inequality (6) holds for any system.

The other possible case is that

$$\Delta r > c\,\Delta t. \tag{10}$$

There can then be no motion representable by the world-vector connecting the two world-points, for the velocity in the reference system considered would have to exceed that of light, and that is impossible. Hence it is also impossible to take the world-vector as a time-axis by a transformation into any other reference system. If the inequality (10) holds for only a single reference system, $(\Delta s)^2$ is *positive* in *every* reference system. The world-vector can then be used only as a *spatial axis*. If we suppose its direction to be that of the $x$-axis, the $y$- and $z$- as well as the time-coordinates are the same for both world-points, and therefore, when the inequality (10) holds good, the two world-points can be made *simultaneous*.

In the two opposite cases of the inequalities (6) and (10), the world-vector connecting the two world-points has different properties. In the first case it can be taken only as a time-axis, in the second case only as a space-axis. In the first instance we speak therefore of a *temporal*, and in the second instance of a *spatial, world-vector*. In the first case the world-points connected by the

## § 182. Fizeau's Experiment.

In the formula which, according to the theory of relativity, holds for the addition of two similarly directed world-vector can be made to coincide in space, and in the second case they can be made simultaneous; the reverse is possible in neither case.

The difference between temporal and spatial world-vectors can more readily be appreciated if, instead of the imaginary time-coordinate ($ict$), we introduce a *real coordinate* equal to $ct$ and denoted by $u$. For simplicity's sake we shall at first consider only a single spatial coordinate, and we construct a rectangular coordinate system in the $x$-$u$-plane. The Lorentz-transformation then appears graphically (Fig. 133) as a transition to the new axes $Ox'$ and $Ou'$, where the angle included between the new and the old axes is $\psi$ and the angle between the new axes themselves is $(90° - 2\psi)$; the tangent of the angle $\psi$ is equal to the quotient of the relative velocity of the two reference systems and the velocity of light. The magnitude of the units of length and time in the new system can be very simply derived from this graphical representation, as we shall show (without, however, giving the proof here). To do this we construct the two rectangular *hyperbolas*

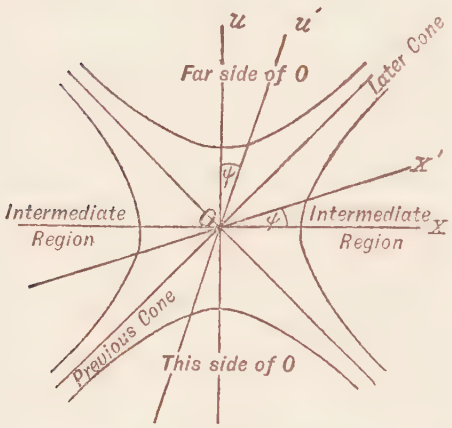



FIG. 133.

$$x^2 - u^2 = \pm 1.$$

The intercepts of the hyperbolas on the $x'$- and $u'$-axes then represent the new units of length and time (the latter multiplied by $c$), referred to the old system. The *pair of asymptotes*

$$x^2 - u^2 = 0$$

represents the propagation of a light wave. If we imagine a $y$-axis to be introduced, the pair of asymptotes is replaced by a double cone, and the further addition of a $z$-axis leads to a double conical super-surface which is called a *light-cone*. The light-cone consists of a *previous cone* and a *later cone*. The former ($u < 0$) represents the propagation of a light-wave which, at the time $t = 0$, arrives at the point $(x = y = z = 0)$; the latter represents the propagation of a wave which, at the time $t = 0$, departs from the point $(x = y = z = 0)$.— The light-cone divides the four-dimensional $x$-$y$-$z$-$u$-manifold (which, however, does not possess the invariant properties of the Minkowski-world) into *three zones*. The first zone lies on the *far side* of $O$, and, in the plane of the figure, is bounded by the two asymptotic branches for which $u$ is positive. To every world-point in this zone a temporal vector can be drawn leading to a world-point which is *later* than the origin. The second zone lies on the *near side* of $O$, and, in the plane of the figure, is bounded by the two asymptotic branches for which $u$ is negative. To every world-point in this zone a temporal vector can be drawn leading to a world-point which is *earlier* than $O$. The third zone, consisting of the rest of the manifold, is the double-branched *intermediate zone*. Spatial vectors can be drawn from the origin to the world-points in it, so that the points in the intermediate zone can be made simultaneous with the origin.

velocities [§ 181 eqn. (4)], we shall take $v'$ as being the *velocity of light in a medium* with a *refractive index*, $n$; thus, let $v'$ be equal to $c/n$. We shall suppose the quantity $v$ in § 181 eqn. (4) to denote the velocity with which the medium itself is moving, as a *flowing liquid* or gas. Then, in accordance with § 181 eqn. (4),

$$(1) \qquad u = \frac{v + \frac{c}{n}}{1 + \frac{v}{c\,n}}.$$

We multiply numerator and denominator by the expression $(1 - v/c\,n)$, neglecting the very small quantity $(v/c\,n)^2$ in comparison with unity. Thus we find

$$(2) \quad u = \left(v + \frac{c}{n}\right)\left(1 - \frac{v}{c\,n}\right) = v + \frac{c}{n} - \frac{v^2}{c\,n} - \frac{v}{n^2} = v - \frac{v}{n^2} + \frac{c}{n}\left(1 - \frac{v^2}{c^2}\right).$$

If again we neglect the very small fraction $v^2/c^2$ as compared with unity, we obtain the approximation formula

$$(3) \qquad u = v\left(1 - \frac{1}{n^2}\right) + \frac{c}{n}.$$

Now, if the propagation of light be unaffected by the flow of the liquid or gas, $u$ must be equal to $c/n$. If, on the other hand, the light-waves be simply carried along by the fluid, $u$ must be equal to the sum of $v$ and $c/n$. Eqn. (3) shows, however, that neither is the case. The *light-waves* are, indeed, affected by the *motion of the fluid*, but not to the full extent of its velocity. The extent to which they partake in that motion is given by a *fraction* $\alpha\, v$, the magnitude of which depends on the refractive index of the fluid; for, according to eqn. (3), the value of the quantity $\alpha$, which is known as the *dragging coefficient*, is given by

$$(4) \qquad \alpha = 1 - \frac{1}{n^2}.$$

As long ago as 1851, a famous experiment performed by *Fizeau* demonstrated that light-waves partake in the flow of a liquid with only a certain fraction of the velocity of flow, and he derived empirically the value given by eqn. (4) for

this fraction.[1] Whereas the interpretation of Fizeau's remarkable experimental result had previously demanded the most complicated hypotheses, his experiment finds a simple explanation in Einstein's Principle of Relativity.

## § 183. Relativistic Dynamics.

According to the theory of relativity, an absolute significance attaches to neither space nor time, but only to the Minkowski-world as an intermingling of the two. In relativistic physics, therefore, the relations between spatial vectors (as known to classical physics) must be replaced by such relations between *world-vectors* that the classical relations can be derived, as special cases, from them.

The theory of relativity is faced with a fundamental problem in the derivation of a more general law to replace *Newton's Second Law of Motion*, which must itself result as a *special case* of that more general law, holding with great approximation for motions in which the velocity is small as compared with the velocity of light.

If, now, there be a privileged system for which the more general relation becomes *identical* with the relation of classical mechanics, this system can evidently only be the one in which the particle considered is *at rest*. We shall define this system briefly as the *proper system* of the particle,[1] and the time corresponding to this system as the *proper time*. Denoting the latter by $\tau$, we have, by § 178 eqn. (4) (since the particle is at rest in the proper system),

$$ds = i\,c\,d\tau\,. \tag{1}$$

Our previous considerations have shown that the four directional cosines of a world-line-element, which we denoted by $\alpha_1$, $\alpha_2$, $\alpha_3$, $\alpha_4$, form the components of a four-dimensional vector.[2] We shall denote by $q_x$, $q_y$, $q_z$, $i\,q_t$ the components of the world-vector obtained by multiplying the directional

[1] Since the refractive index differs very little from unity in the case of gases an appreciable dragging effect is not to be observed with them. The quantity $\alpha$ is also known as *Fresnel's* dragging coefficient, for Fresnel, on the basis of a theoretical investigation of the aberration of light, had come to the conclusion that only a portion of the light æther partakes in the earth's motion. This portion was supposed to be that which existed in matter, and it was assumed to partake only to the extent of a fraction $\alpha$ of the velocity, the value found for $\alpha$ by Fresnel being the same as that given in eqn. (4).

[1] When the motion is not uniform and rectilinear, we can, of course, only speak in general of a proper system for a definite small length of path.

[2] For, in three-dimensional coordinates also, the directional cosines form the components of a vector.

cosines by $ic$. Furthermore, as in § 180, we shall put, for brevity,

$$\frac{1}{\sqrt{1-\frac{v^2}{c^2}}} = \beta . \tag{2}$$

Then, according to § 178 eqn. (6),

$$q_x = \beta v_x , \tag{3}$$

and

$$q_t = \beta c . \tag{4}$$

For the *magnitude* of the world-vector in question, which we define as being the *Minkowski-velocity*, we find from eqns. (3) and (4) the relation

$$q^2 = q_x^2 + q_y^2 + q_z^2 - q_t^2 = \beta^2 (v^2 - c^2) . \tag{5}$$

Hence, replacing $\beta$ by its value as given in eqn. (2), we obtain

$$q^2 = -c^2 . \tag{6}$$

If, in accordance with our earlier assumption, we suppose that Newton's Second Law of Motion in its classical form holds exactly for the *proper system*, we obtain for this system three equations of the form

$$K_x = \frac{d}{d\tau}\left(\mu v_x\right) , \tag{7}$$

*etc.*, where $\mu$ denotes the mass, and $K_x$ the $x$-component of the acting force.

This relation suggests that we might *define* a *world-vector*, which we shall call the *Minkowski-force*, by the equations

$$\left.\begin{aligned} P_x &= \frac{d}{d\tau}\left(\mu q_x\right) , \\ P_t &= \frac{d}{d\tau}\left(\mu q_t\right) . \end{aligned}\right\} \tag{8}$$

Now, if $A$ be an arbitrary quantity, we have by eqn. (1) and § 178 eqn. (5)

$$\frac{dA}{d\tau} = ic \cdot \frac{dA}{dt} \cdot \frac{dt}{ds} = \beta \frac{dA}{dt} . \tag{9}$$

Hence, on substituting for $q_x$ and $q_t$ the values given in eqns. (3) and (4), we find

$$\left.\begin{aligned} P_x &= \beta \frac{d}{dt}\left(\beta \mu v_x\right) , \\ P_t &= \beta c \frac{d}{dt}\left(\beta \mu\right) . \end{aligned}\right\} \tag{10}$$

For the proper system (for which $\beta$ is equal to unity), the first of eqns. (10) actually becomes identical with eqn. (7), and thus we see that the spatial components of the Minkowski force coincide in the proper system with the components of the "Newtonian" force. Consequently, the first of eqns. (10) and the two analogous expressions for $P_y$ and $P_z$ represent a *generalization of Newton's Second Law of Motion*, in virtue of which the *force* is *equal to the time-derivative of the momentum*, as measured by the product of mass and velocity. In this generalization we have, on the one hand, to put

$$P_x = \beta K_x , \tag{11}$$

and, on the other hand, we must relinquish the classical conception of the constancy of mass, assuming, instead, a *dependency of the so-called momental mass on the velocity* according to the relation

$$m = \beta \mu = \frac{\mu}{\sqrt{1 - \frac{v^2}{c^2}}} . \tag{12}$$

In these circumstances, the first of eqns. (10) is transformed, quite generally, into the relation which expresses Newton's Second Law of Motion:

$$\mathbf{K} = \frac{d}{dt}\left(m \mathbf{v}\right) . \tag{13}$$

When $v = 0$, $m$ becomes equal to $\mu$, and $\mu$ is therefore known as the *stationary mass*. The momental mass is always *larger* than the stationary mass. For motions the velocity of which is small as compared with that of light, however, the difference between momental and stationary mass is quite inappreciable.

If, in accordance with eqn. (12), we take into account the dependence of the mass on the velocity, and thus on the time, we can write eqn. (13) in the form

$$\mathbf{K} = m \frac{d\mathbf{v}}{dt} + \mathbf{v} \mu \frac{d\beta}{dt} , \tag{14}$$

where

$$\frac{d\beta}{dt} = \frac{\frac{v}{c^2}}{\left(1 - \frac{v^2}{c^2}\right)^{\frac{3}{2}}} \frac{dv}{dt} = \beta^3 \frac{v}{c^2} \frac{dv}{dt} . \tag{15}$$

We shall now suppose the force to be resolved into two

mutually perpendicular components in the plane determined by the directions of the force and the velocity. One of the two components shall have the direction of the motion, and therefore the components may be distinguished as *longitudinal* and *transversal*. We shall denote the former by the subscript 1, and the latter by the subscript 2. If we resolve the velocity in a similar manner, $v_1$ is equal to $v$ and $v_2$ equal to zero. It follows from eqn. (14) that the acceleration (determined by the differential coefficient $d\mathbf{v}/dt$) must lie in the plane determined by the vectors $\mathbf{K}$ and $\mathbf{v}$; let its two components be $b_1$ and $b_2$. According to the generally valid and purely vectorial relation expressed in § 3 eqn. (30), the longitudinal component of the time-derivative of the velocity-vector is then equal to the time-derivative of the magnitude of the velocity; hence

$$b_1 = \frac{dv}{dt}. \tag{16}$$

On resolving eqn. (14), therefore, into a longitudinal and a transversal partial equation, we obtain [bearing in mind eqn. (12)]

$$K_1 = \beta\mu b_1 + \beta^3 \frac{v^2}{c^2}\mu b_1 = \beta\mu b_1\left(1 + \frac{v^2}{c^2}\beta^2\right).$$

But

$$1 + \frac{v^2}{c^2}\beta^2 = 1 + \frac{v^2}{c^2\left(1 - \frac{v^2}{c^2}\right)} = 1 + \frac{v^2}{c^2 - v^2} = \frac{c^2}{c^2 - v^2} = \beta^2.$$

Hence

$$\frac{K_1}{b_1} = \beta^3\mu = \beta^2 m; \tag{17}$$

whilst it follows from eqn. (14), since $v_2$ is equal to 0, that

$$\frac{K_2}{b_2} = \beta\mu = m. \tag{18}$$

Whereas in classical mechanics, therefore, both $K_1/b_1$ and $K_2/b_2$ are equal to the mass, in relativistic mechanics this is only the case for the transversal component.[3] Since $\beta$ must invariably be larger than unity, we have always

$$\frac{K_1}{b_1} > \frac{K_2}{b_2},$$

[3] The quotient $K_1/b_1$ was formerly known as the "longitudinal mass," and the quotient $K_2/b_2$ as the "transversal mass." Nowadays "mass" is generally understood to mean "momental mass."

or

$$\frac{b_2}{b_1} > \frac{K_2}{K_1}.$$

Thus, in general, *force* and *acceleration* are *not similarly directed*, but the direction of the *acceleration* always includes a *larger angle* with the direction of motion than the direction of the force.

In previous sections reference has been made on several occasions to the *experimental confirmations in atomic physics* of the *mass formula* of the relativity theory. Thus, it was mentioned in § 88 that measurements of the *magnetic deflection of β-rays* have demonstrated the actual validity of the assumption (demanded by the formula) of the decrease in the specific charge with increasing velocity.

### § 184. The Inertial Mass of Energy.

In order to form the *scalar product* of the *world-vectors* representing the *Minkowski-force* and the *Minkowski-velocity*, we shall first write § 183 eqn. (8) in the form

$$P_x = \mu \frac{dq_x}{d\tau} + q_x \frac{d\mu}{d\tau}. \tag{1}$$

Then

$$P_x q_x = \mu \frac{d}{d\tau}\left(\frac{q_x^2}{2}\right) + q_x^2 \frac{d\mu}{d\tau}. \tag{2}$$

On forming the analogous equations for the products $P_y q_y$, $P_z q_z$, and $i^2 P_t q_t$, and then adding them together, we find [bearing in mind the fact that, according to § 183 eqn. (6), $q^2$ is equal to $(-c^2)$, and thus represents a universal constant]

$$P_x q_x + P_y q_y + P_z q_z - P_t q_t = -c^2 \frac{d\mu}{d\tau}. \tag{3}$$

Finally, taking § 183 eqns. (3), (9), and (11) into account, we may write this in the form

$$\beta^2 \mathbf{K}\mathbf{v} - P_t q_t = -\beta c^2 \frac{d\mu}{d\tau}. \tag{4}$$

Now the scalar product of the three-dimensional vectors $\mathbf{K}$ and $\mathbf{v}$ simply represents the *work* performed in unit-time. Moreover, in classical mechanics the proper mass $\mu$ is, of course, regarded as being constant, so that the right-hand side of eqn. (4) vanishes. Hence, provided we neglect the infinitesimal difference of $\beta$ from unity for an arbitrarily

slow motion, the product $P_t\, q_t$ simply represents the *time-derivative of the energy*.

If, therefore, we consider a process in which the *energy* of a body *changes without the performance of external work*, *e.g.*, owing to the absorption or emission of radiation, or to the introduction or abstraction of heat, and since we may assume the motion to be arbitrarily slow, we find, denoting the energy by $E$ and putting $\beta$ equal to unity, that eqn. (4) becomes

$$\frac{dE}{dt} = c^2 \frac{d\mu}{dt}. \tag{5}$$

Thus the theory of relativity leads to the extremely important conclusion that every *change in the energy of a body* involves an *alteration in its mass*, by an amount which is equal to the energy change *divided by the square of the velocity of light*.[1] The *Law of the Inertial Mass of Energy*, expressed in eqn. (5), was derived in 1905 by *Einstein* for a special case, and was subsequently placed on a more general basis by *Planck*.

The theory of relativity, therefore, ascribes mass to all energy, and, conversely, energy to every mass:[2] mass and energy become identical concepts, differing only by a universal proportionality factor. Lavoisier's Law of the Conservation of Mass and the Mayer-Joule Law of the Conservation of Energy are united by the theory of relativity into one single principle, for this theory regards as *constant* only the *sum* of the total energy contained in the universe and the total mass of the universe, multiplied by the square of the velocity of light.

## § 185. World-Tensors.

In complete analogy to the relations holding for a three-dimensional system (*vide* § 26), we define 16 quantities as the components of a *world-tensor* when, for a change of the four-dimensional coordinate system, they are *transformed* in

[1] For a more exact, yet simple, proof of the Law of the Inertia of Energy, *cf.* P. Frank, "*Ein einfacher Beweis für die Trägheit der Energie*," in the scientific journal "*Lotos*" (Prag), 70 (1922), p. 301.

[2] The energy associated with every mass, obtained by multiplying the proper mass by $c^2$, is known as the *proper energy* of the mass. As compared with the proper energy, the so-called kinetic energy (which naturally has only a relative significance) is in general vanishingly small. Even the kinetic energy of a discharged rifle bullet would only represent about 1 part in $10^{12}$ of its proper energy.

the same way as the *products formed by multiplying together the components of two world-vectors in pairs.* If, in particular, the components remain unaltered for an interchange of indices, so that

$$t_{hk} = t_{kh} \tag{1}$$

(where both $h$ and $k$ may be any one of the four numbers 1, 2, 3, 4), the world-tensor is said to be *symmetrical.* If, on the other hand, the sign is reversed for an interchange of the indices, so that

$$t_{hk} = -t_{kh}, \tag{2}$$

the world-tensor is defined as being *skew-symmetrical.* In that case the tensor components of the first kind (*i.e.*, those in which the indices are identical) must be equal and opposite to themselves, which is only possible if they vanish. Hence, for a skew-symmetrical tensor,

$$t_{hh} = 0. \tag{3}$$

Thus, whereas *ten* quantities are necessary for the determination of a symmetrical world-tensor, only *six* are required to define a skew-symmetrical tensor, because four of the components vanish, and in value the other twelve are alike by pairs.

The conception of a skew-symmetrical tensor is, of course, in no way peculiar to four-dimensional geometry alone. In three-dimensional geometry, skew-symmetrical tensors are simply the "supplementary auxiliary vectors" which represent directed plane-magnitudes (*vide* § 3) and as "axial" vectors, are contrasted with the so-called "polar" vectors. Quite generally, in a geometry of an arbitrary number of dimensions, the number of the components of a vector is equal to the number of coordinate *axes*, whereas the number of the components of a skew-symmetrical tensor is equal to the number of coordinate *planes.* In three-dimensional geometry both numbers are the same, and to that fact is due the adoption of the careless form of expression which erroneously describes the skew-symmetrical tensors of three-dimensional geometry as axial vectors.

Now we may regard as the components of a spatial axial vector the three out of the six components of a skew-symmetrical world-tensor which do not contain the time index. On the other hand, the three components containing the time index may be regarded as the components, multiplied by $i$, of a spatial polar vector. Hence we can define a

spatial axial vector ($\mathfrak{C}$) and a spatial polar vector ($\mathbf{D}$) by means of the equations

$$(4)\quad C_x = t_{23} = -t_{32},\quad C_y = t_{31} = -t_{13},\quad C_z = t_{12} = -t_{21};$$

and

$$(5)\quad \begin{cases} i\,D_x = t_{41} = -t_{14}, & i\,D_y = t_{42} = -t_{24}, \\ i\,D_z = t_{43} = -t_{34}. \end{cases}$$

The *vector divergence* of any world-tensor can be found quite generally according to the scheme

$$(6)\quad q_1 = \frac{\partial t_{11}}{\partial x_1} + \frac{\partial t_{12}}{\partial x_2} + \frac{\partial t_{13}}{\partial x_3} + \frac{\partial t_{14}}{\partial x_4}$$

[*cf.* the analogous § 36 eqns. (4)]. We shall put

$$q_1 = q_x,\quad q_2 = q_y,\quad q_3 = q_z,\quad q_4 = i\,q_t,$$

and regard the three components $q_x$, $q_y$, $q_z$ as those of a spatial vector $\mathfrak{q}$. Bearing in mind that $x_4$ is equal to $i\,c\,t$, we find for the vector divergence of the skew-symmetrical world-tensor considered

$$q_x = \frac{\partial C_z}{\partial y} - \frac{\partial C_y}{\partial z} - \frac{1}{c}\frac{\partial D_x}{\partial t},$$

and

$$q_t = \frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z}.$$

Hence

$$(7)\quad \mathfrak{q} = \text{rot}\ \mathfrak{C} - \frac{1}{c}\frac{\partial \mathbf{D}}{\partial t},$$

and

$$(8)\quad q_t = \text{div}\ \mathbf{D}.$$

The generalized operation of *forming a rotation* can be readily found from the formula

$$(9)\quad B_{hk} = \frac{\partial A_k}{\partial x_h} - \frac{\partial A_h}{\partial x_k}.$$

According to § 32, the partial derivatives of a vector with respect to the coordinates represent the components of a tensor, and therefore the quantities $B_{hk}$ determined by eqn. (9) are also the components of a world-tensor. This, as we can recognize at once, must be a *skew-symmetrical world-tensor* which is thus derived as the four-dimensional *rotation* of any world-vector.

On partial differentiation of eqn. (9) with respect to an arbitrary coordinate, we obtain

$$\frac{\partial B_{hk}}{\partial x_r} = \frac{\partial^2 A_k}{\partial x_h \partial x_r} - \frac{\partial^2 A_h}{\partial x_k \partial x_r} . \tag{10}$$

Cyclic interchanges of the three indices $h, k, r$ give us the two analogous equations

$$\left\{ \begin{aligned} \frac{\partial B_{kr}}{\partial x_h} &= \frac{\partial^2 A_r}{\partial x_k \partial x_h} - \frac{\partial^2 A_k}{\partial x_r \partial x_h} , \\ \frac{\partial B_{rh}}{\partial x_k} &= \frac{\partial^2 A_h}{\partial x_r \partial x_k} - \frac{\partial^2 A_r}{\partial x_h \partial x_k} . \end{aligned} \right. \tag{11}$$

Adding eqns. (10) and (11) together, we find

$$\frac{\partial B_{hk}}{\partial x_r} + \frac{\partial B_{kr}}{\partial x_h} + \frac{\partial B_{rh}}{\partial x_k} = 0 . \tag{12}$$

This relation is characteristic of every tensor which, in its turn, forms the rotation of a vector.[1]

If we wish to write down individually the equations comprised in eqn. (12),[2] we have to replace $h, k, r$ by 1, 2, 3, and then effect a cyclic interchange of the four indices 1, 2, 3, 4. We find

$$\left\{ \begin{aligned} \frac{\partial B_{12}}{\partial x_3} + \frac{\partial B_{23}}{\partial x_1} + \frac{\partial B_{31}}{\partial x_2} &= 0 , \\ \frac{\partial B_{23}}{\partial x_4} + \frac{\partial B_{34}}{\partial x_2} + \frac{\partial B_{42}}{\partial x_3} &= 0 , \end{aligned} \right. \tag{13}$$

*etc.*

As before, let us now denote the three spatial components of the skew-symmetrical tensor, represented by the rotation, by $C_x$, $C_y$, $C_z$, and the other three by $i D_x$, $i D_y$, $i D_z$, in accordance with eqns. (4) and (5). Eqn. (13) becomes then

$$\left\{ \begin{aligned} \frac{\partial C_z}{\partial z} + \frac{\partial C_x}{\partial x} + \frac{\partial C_y}{\partial y} &= 0 , \\ -\frac{i}{c}\frac{\partial C_x}{\partial t} - i\frac{\partial D_z}{\partial y} + i\frac{\partial D_y}{\partial z} &= 0 , \end{aligned} \right. \tag{14}$$

*etc.* Thus we may also write

[1] When $r$ is equal either to $h$ or to $k$, one of the three terms in eqn. (12) vanishes.

[2] Eqn. (12) is the four-dimensional generalization of the three-dimensional equation

$$\text{div rot } \mathbf{A} = 0$$

[§ 32 eqn. (24)].

(15) $$\text{rot } \mathbf{D} = -\frac{1}{c}\frac{\partial \mathbf{C}}{\partial t},$$

and

(16) $$\text{div } \mathbf{C} = 0.$$

These two equations apply to every world-tensor which itself forms the rotation of a world-vector, and the components of which can be grouped together in three dimensions to form an axial vector **C** and a polar vector **D**.

### § 186. The Electromagnetic Field Tensor.

Of the four *Maxwell's equations*, two contain only the electric field-strength (**E**) and the magnetic field-strength (**H**), but not the charge or current density. According to § 67 eqns. (1), these two have the form

(1) $$\text{rot } \mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{H}}{\partial t}, \quad \text{div } \mathbf{H} = 0.$$

A comparison of these equations with § 185 eqns. (15) and (16) shows us at once that the components of the magnetic and electric field-strengths represent, in a four-dimensional system, the components of a *skew-symmetrical world-tensor* which, in its turn, must be the *rotation of a world-vector*. This world-vector is called the *four-potential*,[1] and the skew-symmetrical world-tensor derived from it by the operation of rotation is known as the *electromagnetic field tensor*. The components of this tensor are given by the relations

(2) $$F_{23} = -F_{32} = H_x; \quad F_{41} = -F_{14} = i E_x, \text{ etc.}$$

The pair of Maxwell's equations which we have not yet considered, containing the charge and current densities

[1] A simple calculation from Maxwell's equations shows us that the $x$-, $y$-, and $z$-components of the four-potential are identical with the components of Maxwell's vector-potential. The time-component is equal to the so-called *field-potential*, which was introduced into classical electrodynamics by *Lorentz*. Lorentz was led to the concept of the field-potential by the following simple consideration. Denoting Maxwell's vector-potential by **V**, we have

$$\mathbf{H} = \text{rot } \mathbf{V},$$

and, since the order of the operations of forming a rotation and partial differentiation can be reversed,

$$\text{rot}\left(\mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{V}}{\partial t}\right) = 0.$$

Again, according to § 32 eqn. (23), every vector of which the rotation vanishes can be regarded as the gradient of a scalar function of the coordinates. Hence we may put

$$\mathbf{E} + \frac{1}{c}\frac{\partial \mathbf{V}}{\partial t} = -\text{grad } \phi$$

and $\phi$ represents Lorentz's field-potential.

besides the electric and magnetic strengths of field, can be formulated in the following way (according to § 67), if we write the density of the convection current in the form $\varrho\,\mathbf{v}$ :—

$$\frac{4\pi}{c}\varrho\,\mathbf{v} = \operatorname{rot}\mathbf{H} - \frac{1}{c}\frac{\partial \mathbf{E}}{\partial t}, \quad 4\pi\varrho = \operatorname{div}\mathbf{E}. \tag{3}$$

In order to arrive at the four-dimensional significance of these two equations, we shall commence with the concept of a quantity of electricity. Since electrical charge is obviously a quantity which has nothing whatsoever to do with a state of motion, we may reasonably assume that a quantity of electricity will not be affected by a Lorentz-transformation. *Density of charge*, however, will be so affected, because the *volume* is altered by a Lorentz-transformation. For, since a length $x$ is reduced to $x\beta$ by the contraction in the direction of motion,[2] whilst there is no contraction in the $y$- and the $z$-direction, the volume ($V$) itself is not a constant quantity, but

$$\beta V = \text{invariant}, \tag{4}$$

*i.e.*, independent of the reference system. Now the density of charge ($\varrho$) is the quotient of the invariant quantity of electricity and the volume, and therefore

$$\frac{\varrho}{\beta} = \text{invariant}. \tag{5}$$

Let us now multiply the components of the Minkowski-velocity [§ 183 eqns. (3) and (4)] by this scalar (which is independent of the reference system), and also by the universal constant $4\pi/c$. In this way we obtain the components of a world-vector, which is known as the *four-current*. Its components have the values

$$s_x = \frac{4\pi}{c}\varrho\,v_x, \quad s_t = 4\pi\varrho. \tag{6}$$

Resolution of the four-current into its spatial components and its temporal component therefore separates it into convection current and charge (referred to unit-volume). This resolution, however, is of course only relative, and, like the resolution of the electromagnetic field tensor into its electric and magnetic components, depends only on the point of view of the observer describing it. An absolute

[2] The quantity $\beta$, used for brevity's sake, has here again the same meaning as defined by § 180 eqn. (2).

significance, independent of the observer's point of view, attaches only to the four-current and the electromagnetic field tensor.

On substituting the values given by eqns. (6) in eqns. (3), we recognize complete agreement with § 185 eqns. (7) and (8), except that **q**, **C**, **D** are now replaced by **s**, **H**, **E** respectively.

In a four-dimensional representation, therefore, Maxwell's equations may be summarized in the extremely simple statement that the *four-current* is the *vector divergence of the electromagnetic field tensor*, which, in its turn, represents the *rotation* of a world-vector. The transition from four to three dimensions is effected by splitting the four-current into convection current and charge, and the field tensor into electric and magnetic strengths of field.

## § 187. The Matter Tensor.

The *proper mass* ($\mu$), like a quantity of electricity, is evidently independent of the reference system, as, according to § 186 eqn. (4), must also be the case for the quotient $\mu/\beta V$, where $V$ denotes the volume. According to § 183 eqn. (12), however, the product $\beta\mu$ represents the momental mass ($m$), and this divided by the volume gives the so-called *mass density*. For simplicity's sake, we shall denote the mass density, like density of charge, by $\varrho$. We have, therefore,

$$\frac{m}{\beta^2 V} = \frac{\varrho}{\beta^2} = \text{invariant}. \tag{1}$$

We derive a world-tensor, which is of great importance in the theory of relativity, by multiplying the products (taken *in pairs*) of the directional cosines of the *world-line-element* representing the motion by the scalar described by eqn. (1). We find, according to § 178 eqns. (6),

$$\left.\begin{aligned} T_{11} &= -\frac{\varrho\, v_x^2}{c^2}, \quad T_{12} = T_{21} = -\frac{\varrho\, v_x\, v_y}{c^2}; \\ T_{14} &= T_{41} = -i\frac{\varrho\, v_x}{c}, \quad T_{44} = \varrho; \end{aligned}\right\} \tag{2}$$

with analogous expressions for the other components.

The purely temporal component of this tensor thus represents nothing else than the *density of matter*; in subsequent discussions this tensor will therefore be briefly termed the *matter tensor*. The space-time-components ($T_{14}$, $T_{24}$, $T_{34}$) have also a simple meaning, for, apart from a universal

factor, they represent the *density of momentum*. When the mechanical velocity is small as compared with the velocity of light, the purely spatial components become small quantities of the second order, and hence, in this case, we find for the sum of the four tensor components of the first kind (important for later considerations)

$$T_{11} + T_{22} + T_{23} + T_{44} = \varrho . \tag{3}$$

## Chapter XXII

# THEORY OF GRAVITATION

### § 188. The Law of Gravitation as a Generalization of the Law of Inertia.

As a result of the Principle of Relativity established by Einstein in 1905, the concept of absolute rest has lost all meaning in physics. In spite of this, however, the concept of absolute motion was, at that time, by no means completely banished by this principle, for the equivalence of two systems moving with respect to each other was, in the first instance, confined to the *special case of a uniform translatory motion*. Although no single privileged reference system can exist, according to the Principle of Relativity, yet there remains a manifold of privileged coordinate systems which are all moving uniformly with respect to each other. These are the fundamental reference systems of mechanics, previously mentioned in § 13, which are also called *inertial systems*. It is only with regard to such systems that the motion of a body appears uniform, when the body is subject only to the influence of its inertia. On the other hand, an observer who refers the process to another system will detect the presence of so-called *inertial forces*, examples of which are *centrifugal force* and the *Coriolis force* (§ 13). The effect of such inertial forces is manifested in the jerk experienced in a railway train on a sudden application of the brakes, or on travelling round a sharp curve, and also in the phenomenon of *Foucault's pendulum experiment*, or in the *oblateness of the earth*, which is referred to the earth's rotation.

Thus we feel tempted to assume that, whereas absolute velocities do not exist, yet there are *absolute accelerations* and *absolute rotations*. In the long run, however, such an assumption would signify a return to the prejudices vanquished by the Principle of Relativity, to the Newtonian hypothesis of an absolute space and an absolute time. We can understand that such a partial and self-contradictory relativity could not permanently satisfy theoretical physicists, least of all the creator of the theory of relativity himself. An ingeniously simple consideration brought

*Einstein* on to the right path to the solution of the problem, although he had still enormous difficulties to surmount before reaching his destination.

In order to understand Einstein's fundamental ideas, we shall imagine a spatial *inertial system* $(x, y, z)$, with reference to which an observer $(A)$ is at rest, and we shall suppose this observer to describe in terms of this system the physical processes that he observes. A particle which is moving, say, in the $z$-direction solely under the influence of its *inertia* will appear to observer $A$ to be moving *uniformly and rectilinearly*; and this will be the case for any body moving with any velocity in any direction, provided no external forces act upon it.

Let us now consider a second spatial coordinate system $(x', y', z')$ arranged in such a way that the two $z$-axes coincide, whilst the two $x$- and the two $y$-axes constantly remain parallel; and we shall suppose the origin of the second system to be moving with a *constant acceleration* $\gamma$ in the direction of the common positive $z$-axis away from the origin of the first system. We shall imagine a second observer $(B)$ to refer to this second system the processes observed by him, including the motion of the previously mentioned particle. To observer $B$ this motion will, of course, appear to be *uniformly accelerated*, the acceleration being of magnitude $\gamma$ and in the direction *opposite* to that of the positive $z$-axis. Let us next consider various bodies of different mass and different chemical nature, bodies composed, say, one of lead, one of stone, one of cork, and so on; they may be moving in any direction and with arbitrary velocities, but only under the influence of their own inertia. As regards observer $B$, all these bodies will obviously exhibit the *same acceleration* $(-\gamma)$, *independently of their mass and chemical nature*. Provided he knows nothing of observer $A$, and of his own acceleration with respect to $A$, he will obviously interpret this result as *indicating* that he is situated in a *homogeneous gravitational field*. For it has been known to physicists since the time of Galileo that it is characteristic of such a field for all bodies situated in it to experience the same acceleration, independently of all their properties.

In describing the physical processes which take place in the neighbourhood of the two observers, it is thus quite immaterial whether we take the point of view of observer $A$, who refers the processes to his own coordinate system, or of observer $B$, who makes use of the other coordinate system and, in addition, assumes the presence of a homogeneous

gravitational field. Physical processes offer no possibility of deciding between the two points of view, which must therefore be regarded as being equally justified.[1] Recognition of this fact, which is generally known as *Einstein's principle of equivalence*, suggests that the problem of gravitation must obviously be closely linked with the problem of general relativity.

Guided in the right direction by the equivalence principle, we can now form some idea of the fundamental notions of the general theory of relativity. Like gravitational forces, *inertial forces* of every kind, being characteristic of the behaviour of a moving body at a definite point in space and at a definite time, are not a consequence of absolute accelerations, but are *determined by the space-time distribution of matter in the universe*. Conversely, however, this also opens up the possibility of a conception, according to which the *law of gravitation* appears as nothing else than the *generalization of Galileo's Law of Inertia demanded by a truly general principle of relativity.*

How, then, is a generalization of this law possible ? The Law of Inertia teaches us that every body, upon which no external forces are acting, moves with constant velocity in a straight path, so that not only is the path thus rectilinear, but also the three spatial coordinates are linear functions of the time. Translated from three into four dimensions, the Law of Inertia therefore states that the *world-line* describing a *motion unattended by the application of force* must be a *straight line*. Now, is this statement capable of further generalization ? As a matter of fact it is, on the basis of a *generalization of the geometrical foundations of physics*.

In order to understand this, we shall, for simplicity's sake, first consider two-dimensional geometry. There can hardly be any doubt as to when a line drawn on a plane sheet of paper is to be regarded as a straight line. A *two-dimensional geometry*, however, is just as possible *on a spherical surface* or, indeed, *on any curved surface whatsoever*, as on a plane ; but the question as to when a line drawn on

[1] The example given by Einstein himself, of a physicist enclosed in a large chest somewhere in space, is particularly illuminating. The chest is suspended by a rope which passes over a pulley, and the other end of the rope is pulled downwards with constant acceleration. If he knows nothing of the world outside the chest, and nothing about the pulley and the rope, the physicist inside the chest will naturally observe the presence of a gravitational field (which, if we imagine the chest to be on the earth instead of in space, will be superimposed on the gravitational field of the earth).

such a surface is to be regarded as a straight line is not so easily answered. Let us consider two points on a globe, which lie on the same parallel of latitude, but differ widely in longitude. Since we are dealing with two-dimensional geometry, we are, of course, unable to leave the surface of the globe. Now what do we mean by the straight line connecting the two points on the surface of the globe ? Is it the portion of the parallel of latitude between them ? If this question is to have any meaning at all, we must obviously first replace the vague conception of a " straight line " by a more general and well-defined one. Such a conception is already long known in geometry, *viz.*, that of a *geodesic line*, which is *defined* as being the *shortest line joining two points on the surface*. On the basis of this definition (to illustrate by an example) we find, by means of a geometrical investigation, that the segment of the parallel of latitude between the two points in the above consideration in no way represents a geodesic line. We do, however, obtain such a line by constructing a great circle on the sphere to pass through both points.[2] In this lies the explanation of the apparently paradoxical fact that an aviator, who wishes to fly by the shortest route from London to Australia, should go *viâ* Petrograd.

Now we can clearly conceive a curved surface as a two-dimensional manifold " embedded " in one of three dimensions. On the other hand, we cannot form any conception of a curved space, for this in itself is of three dimensions. The limitations of the human imagination, however, offer no barrier to an abstract mathematical treatment. After *Gauss*, in 1827, had developed a *general theory of surfaces*, in 1854 the great mathematician, *Riemann*, showed how our usual *three-dimensional solid geometry* can also be regarded as a *special case* of more general three-dimensional geometry, in which space is " curved " just as a surface may be in two-dimensional geometry. Moreover, as Riemann himself recognized, what holds for three-dimensional geometry must also apply to a geometry of *any number of dimensions :* it must therefore also hold for the four-dimensional manifold which, linking time and space together, is known as the *Minkowski-world*. When we base our mathematical treatment of it upon a general instead of a special geometry, we must regard the Minkowski-world also as being *curved*,

[2] We can temporarily transfer the pole to one of the points, and choose from the manifold of meridians passing through it the meridian which passes through the second point.

whilst the degree of curvature may vary, of course, continuously from place to place.

We owe to Einstein the ingenious concept of the curved Minkowski-world and of connecting the curvature with the space-time distribution of matter. Einstein's *generalized Law of Inertia*, which represents at the same time the new *Law of Gravitation*, simply states that every particle always moves in such a way that its *world-line is a geodesic in the curved Minkowski-world*, the *curvature* being *caused by matter*.[3]

This conception of gravitation leads directly to a very important conclusion. Let us suppose a particle with any velocity to be situated at any position in space at any time. We shall consider its mass, however, to be only a "test mass," *i.e.*, to be so small that the field of matter and thus the curvature of the Minkowski-world are not affected by it (just as in electrostatics we consider proof charges which have no influence on the electric field). According to our previous remarks, the motion of the test mass will be represented by the geodesic passing through the world-point considered in the direction corresponding to that of the given velocity. Now let us suppose a test body to be situated at the same place again, at the same time, and to possess the same velocity, but with a *mass half as large* as was originally the case. The motion will be represented by *precisely the same geodesic* as before. On the other hand, the acceleration is completely determined by the world-line for the given reference system, for it depends only on the functional relation connecting the four world-coordinates in the equation of the curve. Hence, if the Law of Gravitation be regarded as a generalization of the Law of Inertia, it follows necessarily that the *acceleration*, which we *interpret as a gravitational effect*, must be *independent of the mass*.

This result can be expressed in another way. It is usual to define the *inertial mass* of a body as the constant ratio which exists between an arbitrary force acting on the body and the acceleration thereby produced. (We shall not take account of the relativity correction here.) We arrive thus at a unit of mass, the magnitude of which depends on the units of force and acceleration employed. On the other hand, we can also determine the mass of bodies by measuring the forces of attraction which they experience in a gravitational

[3] In this connection it is assumed, of course, that *no external forces* are acting on the particle. Under this heading we include electromagnetic forces, but, in view of the now conceptions, not so-called gravitational forces.

field as compared with the force exerted on the above unit-mass in the same field. The number representing this ratio then determines the *gravitational mass* of the body. If, now, the acceleration in the gravitational field is actually independent of the inertial mass, then, according to Newton's Second Law of Motion, the force of gravitation must be exactly proportional to the inertial mass. Hence an inertial mass of twice the previous magnitude must be associated with twice as large a value of the gravitational force, and thus with twice the *gravitational* mass, and so on. In general, *gravitational and inertial mass* must be *identical*.

This conclusion already found definite expression in Galileo's observation that all bodies, whatever their mass, fall at the same rate in empty space. In classical mechanics it was taken for granted that the inertial mass must be proportional to the weight of a body. The correctness of this assumption was confirmed in detail by accurate measurements carried out by *Eötvös* in 1890. These measurements were based on the fact that the weight of a body represents the resultant of two forces, one of which is the attraction of the earth, and the other the centrifugal force due to the rotation of the earth. Whereas the former of these is determined by the gravitational mass of the body, the latter depends on its inertial mass. The experiments of Eötvös were carried out with extraordinary precision, and they showed that, even when the masses differed widely in magnitude and in chemical constitution, no difference whatsoever could be detected in the direction of the resultant force. The law of the identity of gravitational and inertial mass appears to be completely established by these measurements. Whereas, however, this law was regarded as self-evident prior to Einstein's theory, and physicists gave it no further thought, it was Einstein's work that first revealed the profound reason for this remarkable law.

The new Law of Gravitation is, of course, still incomplete, so long as we do not know the relation expressing the way in which the *Minkowski-world* is *curved by matter*. In order to reach a solution of this question, let us imagine a *dual system of sets of curves* on an arbitrarily curved surface. We shall distinguish these curves as $x$- and $y$-curves, and no two $x$-curves nor any two $y$-curves shall intersect (Fig. 134). We can then suppose each of the successive $x$-curves to be denoted by consecutive numbers, and similarly for the $y$-curves. We can also imagine, however, nine intermediate curves to be drawn between, say, the curves ($x = 6$) and

($x = 7$), which likewise shall not intersect each other, and we can assign to these curves the numbers 6·1, 6·2, . . . 6·9. Similarly, we can insert nine new curves between the curves ($x = 6·1$) and ($x = 6·2$), denoting them by the numbers 6·11, 6·12, . . . 6·19. In this way we can finally assign to every point on the surface two numerical values, an $x$- and a $y$-value, and regard these in the broader sense of the word as the coordinates of the point in question. These values are called the *Gaussian coordinates* of the point. The *Cartesian* coordinate systems used in plane geometry represent, of course, only *special cases* of Gaussian systems.

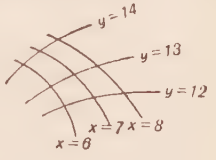


FIG. 134.

Having thus established a Gaussian coordinate system, let us consider three neighbouring points in the surface, which make up a *triangle*. In the particular case of plane coordinate geometry, the coordinates of the three points would completely determine the lengths of the sides, the angles, and the area of the triangle. In the geometry of an arbitrarily curved surface, however, this is not the case. The statement of the coordinates does *not* then suffice for the determination of the lengths of the sides, the angles, and the area of the triangle. In addition to this, we require the values of three further quantities for the region considered, which we may denote by $g_{xx}$, $g_{xy}$, $g_{yy}$; and, as was recognized by Gauss, these three quantities represent the *components of a tensor*. Because this tensor determines the measure relations in any particular region, it is called the *metric fundamental tensor*. Its value must be unequivocally *stated* for the place at which we supposed the small triangle to be constructed, if we wish to calculate the sides, the angles, and the area of the triangle from the coordinates of the three points. It determines really the kind of geometry,[4] and therefore it enters necessarily into all formulæ of a

[4] It is, of course, well known that the geometry of spherical surfaces is governed by other theorems than those applying to plane, Euclidean, geometry. For example, Euclid's well-known theorem, which states that the sum of the angles in a triangle amounts to two right angles, does not hold in spherical geometry. In point of fact, if we imagine a spherical triangle on the terrestrial globe formed by the northern quadrants of the meridians of 0° and 90° east longitude, and by the corresponding quadrant of the Equator, this triangle contains three right angles, of which the sum is 270°.—The theorems of plane geometry apply, moreover, not only to planes, but also to such surfaces as can be produced by the bending of a plane surface, *e.g.*, cylindrical or conical surfaces.

general theory of surfaces. Conversely, we can also determine the fundamental tensor from measurements on the above-mentioned triangle, without our requiring to know anything as to how the surface is " embedded in space " at the place in question.

Thus a general geometry of surfaces only has any sense and only becomes possible when the fundamental tensor is given for every position on the surface. Now what holds good for a general two-dimensional geometry likewise applies to any general geometry of an arbitrary number of dimensions. The fundamental tensor is symmetrical, and hence in two-dimensional geometry it has three, in three-dimensional six, and in four-dimensional geometry ten components ; in general, of course, it will *vary continuously from place to place*. Every geometrical manifold may therefore also be regarded as the field of its metric fundamental tensor or, in brief, as a *metric field*.

Purely mathematical investigations show that a scalar, which is known as the *Riemann scalar* and is completely independent of the coordinate system, can be derived from the fundamental tensor. This scalar, of course, will also in general vary continuously from place to place. Let us imagine any region to be delimited within the geometrical manifold, and the Riemann scalar to be integrated over it. Then we can calculate the *variation of this integral* on the assumption that, *within* the region, the components of the fundamental tensor may be varied arbitrarily and independently of each other, conditionally upon the variations vanishing at the boundary itself. The process is entirely similar to that of the variation, in mechanics, of a motion between the initial and final positions, on the condition that the actual and the varied motion shall coincide at these positions. If, now, we calculate the variation of the integral as stated, we obtain a tensor *independent of the coordinate system*, which, in our considerations, may be called the *curvature tensor*. (In order to avoid any misunderstanding, however, it should be expressly noted that the title of " curvature tensor " is often used in another sense.) The derivation of this tensor follows along purely mathematical lines, quite independently of all physics.

*Einstein's Law of Gravitation*, which states that the world-line of a particle shall be a geodesic, is completed by the further postulate that, at every point in the Minkowski-world, the *curvature tensor shall be identical with the matter tensor* (§ 187), apart from a *universal propor-*

*tionality factor* determined by the usual physical units of measure.

Since both the geodesic and the curvature tensor are independent of the Gaussian coordinate system used, it follows that Einstein's Law of Gravitation or, in other words, the generalized Law of Inertia, also holds for *any Gaussian coordinate system whatsoever*. As long as the theory of relativity confined itself to Cartesian coordinate systems in its treatment of the Minkowski-world, it could use the displacement of the origin and the rotation of the axes in a four-dimensional coordinate system only to effect transitions between such spatial coordinate systems as are *moving uniformly with respect to each other*; for no four-dimensional coordinate transformations other than *linear* were available. The application of Gaussian coordinate systems opened up an entirely new field. A change, which in space appears as a transition between two coordinate systems *moving quite arbitrarily with regard to each other*, is always represented in a four-dimensional manifold by a transformation from one to some other Gaussian coordinate system.

The generalized Law of Inertia, when in the form of Einstein's Law of Gravitation, must therefore hold in exactly the same way for two different spatial coordinate systems, no matter how these are moving relatively to each other. On a transition from one system to the other, however, the *components* of the fundamental tensor are transformed, and it is just the alterations thus connected with the transition which were interpreted by classical mechanics as *inertial forces*, and these differ according to the choice of reference system.

From this new point of view, the battle that once raged so violently between the Ptolemaic and the Copernican systems of the universe appears quite pointless. If we imagine the earth at rest and the sky with the fixed stars rotating around it, the curvature of the Minkowski-world must yet be such that measurements of the earth's surface will show its oblateness, as *Thirring* [5] has demonstrated on the basis of Einstein's theory of gravitation.

Just as the generalized Law of Inertia is *invariant* in arbitrary four-dimensional coordinate transformations, and thus the equations referring to the components hold for any coordinate system, so also the *equations of the electromagnetic field* (§ 186) can be reduced to a form which holds for any coordinate system. The relation connecting a vector with

[5] H. Thirring, *Physikal. Zeitschr.*, **19** (1918), p. 33.

its rotation, or a skew-symmetrical tensor with its vector divergence, is something quite independent of the coordinate system; but the equations connecting the *components* are, of course, more complicated when based on Riemannian geometry than in Euclidean geometry.

The postulate of relativity thus appears fulfilled in the most complete manner by the new theory which, with the help of Einstein's Law of Gravitation, bases the description of the Minkowski-world on Riemannian geometry. For this reason this theory, which was developed by Einstein in 1915, is known as the *general theory of relativity*, in contrast to the special theory of relativity which represents only a special case.

## § 189. Gaussian Coordinate Systems.

In order to be able to develop the mathematical bases of Einstein's theory of gravitation in an exact manner, we shall first consider *two neighbouring points on an arbitrarily curved surface*. Let the coordinates of these points relatively to an arbitrary Gaussian coordinate system be $x$, $y$ and $(x + dx)$, $(y + dy)$ respectively; and let the *distance* between the two points (measured in the surface) be $ds$. Since we can regard an *infinitesimally small portion of the surface as plane*, we can also suppose a Cartesian coordinate system to be constructed in the small region of the surface containing the two points considered; with reference to this system, let the coordinates of the first point be $\xi$ and $\eta$, and those of the second $(\xi + d\xi)$ and $(\eta - d\eta)$. The quantities $x$ and $y$ can then undoubtedly be regarded as *functions* of the two quantities $\xi$ and $\eta$; and, if we solve the two equations expressing this connection with respect to $\xi$ and $\eta$, we obtain two other equations

$$\xi = \varphi_1(x, y); \qquad \eta = \varphi_2(x, y). \tag{1}$$

Differentiating these, we find

$$\left\{ \begin{aligned} d\xi &= \frac{\partial \varphi_1}{\partial x}\, dx + \frac{\partial \varphi_1}{\partial y}\, dy, \\ d\eta &= \frac{\partial \varphi_2}{\partial x}\, dx + \frac{\partial \varphi_2}{\partial y}\, dy. \end{aligned} \right. \tag{2}$$

In the Cartesian coordinate system, the square of the distance between the two points is given by

$$ds^2 = d\xi^2 + d\eta^2, \tag{3}$$

and hence, in the Gaussian system, we have by eqns. (2)

$$(4)\quad \left\{\begin{aligned} ds^2 = &\left[\left(\frac{\partial\varphi_1}{\partial x}\right)^2 + \left(\frac{\partial\varphi_2}{\partial x}\right)^2\right] dx^2 + \left[\left(\frac{\partial\varphi_1}{\partial y}\right)^2\right. \\ &\left. + \left(\frac{\partial\varphi_2}{\partial y}\right)^2\right] dy^2 + 2\left[\frac{\partial\varphi_1}{\partial x}\cdot\frac{\partial\varphi_1}{\partial y} + \frac{\partial\varphi_2}{\partial x}\cdot\frac{\partial\varphi_2}{\partial y}\right] dx\,dy\,. \end{aligned}\right.$$

We shall denote by $x_h$ either of the two coordinates $x$ or $y$, according as the subscript $h$ is put equal to 1 or 2, and thus we can write eqn. (4) in the form

$$(5)\qquad ds^2 = \sum_h \sum_k g_{hk}\, dx_h\, dx_k \qquad (h = 1, 2\,;\ k = 1, 2)\,;$$

where, by eqn. (4),

$$(6)\qquad g_{hk} = \frac{\partial\varphi_1}{\partial x_h}\cdot\frac{\partial\varphi_1}{\partial x_k} + \frac{\partial\varphi_2}{\partial x_h}\cdot\frac{\partial\varphi_2}{\partial x_k}.$$

Hence

$$(7)\qquad g_{hk} = g_{kh}\,.$$

If, in eqns. (5) to (7), we give the subscripts $h$ and $k$ three, four, or more values, instead of only two, and if, correspondingly, we introduce $\varphi_3$, $\varphi_4$, *etc.*, in addition to $\varphi_1$ and $\varphi_2$, we obtain the fundamental formulæ of a *Riemannian geometry of an arbitrary number of dimensions.* In two-dimensional geometry there are four quantities $g_{hk}$, of which two are the same ; in three-dimensional geometry there are nine quantities, of which three pairs are alike ; and in four-dimensional geometry there are sixteen quantities, including six pairs alike, according to the scheme

$$(8)\qquad \left\{\begin{matrix} g_{11} & g_{12} & g_{13} & g_{14} \\ g_{21} & g_{22} & g_{23} & g_{24} \\ g_{31} & g_{32} & g_{33} & g_{34} \\ g_{41} & g_{42} & g_{43} & g_{44}\,, \end{matrix}\right.$$

where

$$g_{12} = g_{21}, \qquad g_{13} = g_{31},\ \textit{etc.}$$

Eqn. (5) represents the *generalization of the Theorem of Pythagoras in Riemannian geometry.* We see at once from this equation that *Cartesian coordinate systems* represent special cases of Gaussian systems, for orthogonal Cartesian systems are obtained by equating the quantities $g_{hk}$ to unity or zero according as the two subscripts are the same or different. In such systems

$$(9)\qquad g_{hk} = \left\{\begin{matrix} 1\ (h = k) \\ 0\ (h \neq k) \end{matrix}\right.;$$

and, in this special case, scheme (8) assumes the simple form

$$(10) \qquad \left\{ \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \, . \end{array} \right.$$

## § 190. Generalized Conception of Tensors.

In the previous section we considered a Gaussian coordinate system in which the coordinates of two neighbouring points on an arbitrary surface were $x$, $y$ and $(x + dx)$, $(y + dy)$ respectively. We shall now suppose this system to be replaced by a second, likewise Gaussian, coordinate system, with respect to which the coordinates of the points may be denoted by $x'$, $y'$ and $(x' + dx')$, $(y' + dy')$ respectively. Since the new coordinates may be regarded as functions of the old ones, we have

$$(1) \qquad \left\{ \begin{array}{l} dx' = \dfrac{\partial x'}{\partial x} dx + \dfrac{\partial x'}{\partial y} dy \, , \\ dy' = \dfrac{\partial y'}{\partial x} dx + \dfrac{\partial y'}{\partial y} dy \, . \end{array} \right.$$

Let us now imagine any *force*, and let its components relatively to the original coordinate system be $X$ and $Y$, and those relatively to the new one be $X'$ and $Y'$. Then the *work* done by this force along the element of path, the components of which are $dx$ and $dy$, or $dx'$ and $dy'$ respectively, must undoubtedly be independent of the coordinate system; *i.e.*, it must be, as we say, *invariant* as regards coordinate transformations. Thus we must have

$$(2) \qquad X' \, dx' + Y' \, dy' = X \, dx + Y \, dy \, .$$

Introducing in this equation the values given for $dx'$ and $dy'$ by eqns. (1), we find

$$(3) \qquad \left\{ \begin{array}{l} \left( X' \dfrac{\partial x'}{\partial x} + Y' \dfrac{\partial y'}{\partial x} \right) dx + \left( X' \dfrac{\partial x'}{\partial y} + Y' \dfrac{\partial y'}{\partial y} \right) dy \\ \qquad = X \, dx + Y \, dy \, . \end{array} \right.$$

If this relation is to be *generally* fulfilled, the equations

$$(4) \qquad \left\{ \begin{array}{l} X = X' \dfrac{\partial x'}{\partial x} + Y' \dfrac{\partial y'}{\partial x} \, , \\ Y = X' \dfrac{\partial x'}{\partial y} + Y' \dfrac{\partial y'}{\partial y} \end{array} \right.$$

must be satisfied. Since the transition from the one coordinate system to the other is in no way privileged as compared with the reverse transition, the relation

$$X' = X\frac{\partial x}{\partial x'} + Y\frac{\partial y}{\partial x'} \tag{5}$$

must, of course, also hold good.

Let us now pass from the two-dimensional system hitherto considered to a coordinate system of an arbitrary number of dimensions ; then the *transformation formulæ* of eqns. (1) and (5) can be written in the more general form

$$dx'_h = \sum_k \frac{\partial x'_h}{\partial x_k} dx_k , \tag{6}$$

and

$$X'_h = \sum_k X_k \frac{\partial x_k}{\partial x'_h}. \tag{7}$$

On the basis of these equations we *define* a quantity as a *covariant vector* when its components obey the transformation law

$$A'_h = \sum_k A_k \frac{\partial x_k}{\partial x'_h}. \tag{8}$$

On the other hand, we define a quantity as a *contravariant vector* when its components are transformed in accordance with the law

$$A'^h = \sum_k A^k \frac{\partial x'_h}{\partial x_k}. \tag{9}$$

It is customary, as has been done in eqns. (8) and (9), to distinguish *covariant* vectors by writing the *index below*, and *contravariant* vectors by writing it *above*. In Euclidean geometry, of course, no distinction is made between covariant and contravariant vectors, but this is essential in Riemannian geometry.

If $A_h$ be the components of an arbitrary covariant vector, and $B^h$ those of an arbitrary contravariant vector, it follows at once from eqns. (8) and (9) that

$$\sum_h A'_h B'^h = \sum_k A_k B^k,$$

so that

$$\sum_h A_h B^h = \text{invariant}. \tag{10}$$

In our further considerations we shall introduce a *simplified notation.* For this and all subsequent sections we stipulate that *everywhere that the same index occurs twice in a term, the term has to be summed for that index, although the sign of summation is dropped.*

Corresponding to the distinction between covariant and contravariant vectors, we define as a *covariant*, a *contravariant*, or a *mixed tensor of the second rank* a quantity the components of which are transformed in the same way as the products formed by multiplying in pairs the components of two covariant vectors, or of two contravariant vectors, or of one covariant and one contravariant vector. Thus a covariant tensor of the second rank is transformed according to the formula

$$A'_{pr} = A_{hk} \frac{\partial x_h}{\partial x'_p} \cdot \frac{\partial x_k}{\partial x'_r}. \tag{11}$$

(Following the rule given above, the double summation symbol is omitted.)

On the other hand, the transformation of a contravariant tensor of the second rank takes place in accordance with the formula

$$A'^{pr} = A^{hk} \frac{\partial x'_p}{\partial x_h} \cdot \frac{\partial x'_r}{\partial x_k}. \tag{12}$$

It follows from eqns. (11) and (12) that for two arbitrary tensors, one of which is covariant and the other contravariant, the double sum

$$A_{hk} B^{hk} = \text{invariant}. \tag{13}$$

Finally, the components of a mixed tensor follow the transformation law

$$A'^{r}_{p} = A^{k}_{h} \frac{\partial x_h}{\partial x'_p} \cdot \frac{\partial x'_r}{\partial x_k}. \tag{14}$$

Tensors of a rank higher than the second are defined in an analogous manner. Thus

$$A^{r}_{hkp} = B_h C_k D_p E^r \tag{15}$$

would be an example of a *mixed tensor of the fourth rank, etc.* In a four-dimensional system, a tensor of the third rank has in general 64, one of the fourth rank 256, and one of the $z^{th}$ rank $4^z$ components.

The property of invariance, as expressed in eqn. (10), enables us to derive from any mixed tensor of arbitrary rank another *tensor* of a *rank lower by two*, by equating an upper

to a lower index, and then summing for this index. *E.g.*, by means of this operation we can derive a tensor of the second rank from the mixed tensor of the fourth rank defined by eqn. (15), with the help of the relation

$$A_{hk} = A^{p}_{hkp} . \tag{16}$$

In point of fact, if for the moment we replace the summation symbol (otherwise omitted), we have

$$A_{hk} = \sum_{p} A^{p}_{hkp} = B_h C_k \sum_{p} D_p E^p . \tag{17}$$

Since, however, the sum $D_p E^p$ is [according to eqn. (10)] invariant with respect to coordinate transformations, the quantity $A_{hk}$ defined by eqn. (16) is actually a covariant tensor of the second rank.

This operation, by means of which we can thus derive from an arbitrary tensor [1] of rank $z$ another tensor of rank $(z - 2)$, is known as the *contraction of a tensor*. The contraction of a mixed tensor of the second rank leads to a tensor of zero rank, *i.e.*, a *scalar*.

## § 191. The Metric Fundamental Tensor.

From § 190 eqn. (13) we are able at once to draw an important conclusion with respect to the quantities $g_{hk}$ which, when use is made of a Gaussian coordinate system, occur in the expression for the line-element. The quantity $ds^2$, occurring in § 189 eqn. (5), is invariant and thus independent of the coordinate system. According to § 190 eqn. (6), the quantities $dx_h$ and $dx_k$ are the components of two contravariant vectors, and hence the product $dx_h \cdot dx_k$ must be a component of a contravariant tensor of the second rank. On the other hand, however, we have by § 189 eqn. (5)

$$ds^2 = g_{hk}\, dx_h\, dx_k , \tag{1}$$

and therefore, according to § 190 eqn. (13), the quantities $g_{hk}$ must be the components of a *covariant tensor of the second rank*. This tensor is the *metric fundamental tensor*; [1] the geometrical manifold, in which, in general, it varies continuously from place to place, thus appears as the field of this fundamental tensor or, in brief, as the *metric field*.

Some simple calculations enable us to derive from the

[1] For any arbitrary tensor can be reduced to one, the components of which can be regarded as the products of the components of vectors.

[1] It is called "metric" because, in accordance with § 189 eqn. (5), it is the fundamental tensor that determines the measure relations ("metrics").

sixteen components of the metric fundamental tensor sixteen other quantities, which are likewise of importance for subsequent considerations. We can regard the sixteen quantities in scheme (8) of § 189 as the *elements of a determinant*, and form the *minor* of each $g_{hk}$ by striking out the column and row containing the element in question.[2] We obtain the *cofactor* of the element by multiplying the minor by $(-1)^{h+k}$, where $h$ and $k$ are given the values 1, 2, 3, or 4, as the case may be.[3] The *value* ($g$) of the determinant is obtained, in accordance with a well-known theorem, on multiplying all the elements in any row or column by the corresponding cofactors, and then adding together.[4] If we denote the cofactor of $g_{hk}$ by $G_{hk}$, we have

$$g = g_{1k} \cdot G_{1k} = g_{2k} \cdot G_{2k} = g_{3k} \cdot G_{3k} = g_{4k} \cdot G_{4k} ,$$

where, in accordance with our rule of notation, we have to take the sum for the index $k$. Consequently, the double sum

$$g_{hk} G_{hk} = 4\, g ,$$

or,

$$g_{hk} \frac{G_{hk}}{g} = 4 . \tag{2}$$

Since the number 4 is, of course, invariant with respect to coordinate transformations, it follows from § 190 eqn. (13) that the quantities $G_{hk}/g$ must be the components of a contravariant tensor of the second rank. This tensor is known as the *contravariant fundamental tensor*, and its components are denoted by the symbol

$$g^{hk} = \frac{G_{hk}}{g} . \tag{3}$$

On multiplying all the components of the covariant by all the components of the contravariant fundamental tensor, we obtain a mixed tensor of the fourth rank:

$$g^{kr}_{hp} = g_{hp} \cdot g^{kr} . \tag{4}$$

[2] The minor of the element $g_{23}$ is given by

$$G_{23} = \begin{vmatrix} g_{11} & g_{12} & g_{14} \\ g_{31} & g_{32} & g_{34} \\ g_{41} & g_{42} & g_{44} \end{vmatrix} .$$

[3] For the cofactor $G_{23}$ we have $h + k = 5$, and hence its sign is negative: on the other hand, $G_{13}$ or $G_{24}$ is positive.

[4] The values of the cofactors themselves are calculated by the same rule *e.g.*,

$$-G_{23} = g_{11}(g_{22}g_{44} - g_{34}g_{42}) - g_{12}(g_{31}g_{44} - g_{34}g_{41}) + g_{14}(g_{31}g_{42} - g_{32}g_{41}).$$

Contraction of this tensor (by putting $r$ equal to $p$, and then summing for $p$) leads to the so-called *mixed fundamental tensor*. Thus we have

$$g_h^k = g_{hp} \cdot g^{kp} = \frac{g_{hp}\, G_{kp}}{g} . \tag{5}$$

Now when the index $h$ is equal to the index $k$, the numerator of the fraction simply represents the value obtained when the elements in one row of a determinant are multiplied by the corresponding cofactors, and the products added together. This, however, gives the value ($g$) of the determinant, so that

$$g_h^k = 1, \text{ when } k = h.$$

Let us next consider the case in which $k$ differs from $h$. We obtain then the sum $g_{hp} \cdot G_{kp}$ on multiplying the elements in the $h^{th}$ row by the cofactors of the elements in the $k^{th}$ row. The value so obtained, however, is identical with the value of the determinant formed by replacing the elements in the $h^{th}$ row by those in the $k^{th}$ row.[5] We should thus have a determinant in which the $h^{th}$ and $k^{th}$ rows were identical. As may readily be seen, however, the value of a determinant in which any two rows are identical is zero. For the sign of a determinant is reversed when two of its rows are interchanged,[6] and hence, for a determinant with two identical rows, $g$ must be equal to $(-g)$, which is, of course, only possible when $g$ vanishes. Therefore

$$g_{hp}\, G_{kp} = 0, \text{ when } k \neq h.$$

Hence we find quite generally for the mixed fundamental tensor, for which the symbol $\delta_h^k$ is usual,

$$g_h^k = \delta_h^k = \begin{cases} 1 \ (k = h) \\ 0 \ (k \neq h) \end{cases} . \tag{6}$$

Finally, we shall show how, with the help of the fundamental tensors, a contravariant vector can be derived from a given covariant vector, and *vice-versâ*. If $A_h$ be an

[5] *E.g.*, the value of $g_{1p}\, G_{2p}$ is that of the determinant

$$\begin{vmatrix} g_{21} & g_{22} & g_{23} & g_{24} \\ g_{21} & g_{22} & g_{23} & g_{24} \\ g_{31} & g_{32} & g_{33} & g_{34} \\ g_{41} & g_{42} & g_{43} & g_{44} \end{vmatrix} .$$

[6] This well-known theorem follows from the previously quoted rule for the signs of the cofactors.

arbitrary covariant vector, we can form a tensor of the third rank :

$$A_h^{kp} = g^{kp} A_h .$$

On contracting this mixed tensor, *i.e.*, on putting $p$ equal to $h$ and summing for $h$, we obtain a contravariant vector :

$$A^k = g^{hk} A_h . \tag{7}$$

Conversely, we can derive a covariant from a contravariant vector ($B^k$) with the aid of the relation

$$B_p = g_{kp} B^k . \tag{8}$$

If, in particular, we replace $B^k$ by the vector $A^k$ as defined in eqn. (7), we find for $B_p$ the value

$$B_p = g_{kp} g^{hk} A_h = \delta_p^h A_h = A_p ; \tag{9}$$

*i.e.*, the vector $A_h$ belongs in the same way to the vector $A^h$ as the vector $A^h$ does to the vector $A_h$. Hence, instead of speaking of a covariant vector $A_h$ and a contravariant vector $A^h$, we can refer simply to a single vector, of which the *covariant components* are $A_h$, and the *contravariant components* are $A^h$.

In an entirely analogous manner, we can derive a mixed tensor from a covariant tensor of the second rank ($A_{hk}$) by multiplying by $g^{pr}$, and then contracting with respect to the indices $r$ and $k$ :

$$A_h^p = g^{pk} A_{hk} . \tag{10}$$

On a further multiplication by $g^{si}$ and subsequent contraction with respect to the indices $h$ and $i$, we find

$$A^{ps} = g^{pk} g^{sh} A_{hk} . \tag{11}$$

In this way we can derive a mixed and a contravariant tensor from any covariant tensor, and thus we may also speak of the covariant, the contravariant, and the mixed components of one and the same tensor.

## § 192. Geodesic Lines.

In the last section we have formed an idea of the fundamental tensor, and we are now faced with another problem, which is likewise purely mathematical, of finding the *general equation of a geodesic line*, which is defined as being the *shortest line joining* two arbitrary points $A$ and $B$ within a geometrical manifold. A geodesic line must therefore satisfy the condition

$$\delta \int_A^B ds = 0 , \tag{1}$$

where $ds$ denotes an element of the geodesic, and the variation is to be performed in such a way that the varied curve also leads from the point $A$ to the point $B$.[1]

Let us suppose a set of $(n-1)$-dimensional super-surfaces to be constructed in an $n$-dimensional manifold, the equations of these surfaces being of the form

$$\lambda = \varphi(x_1, x_2, x_3, \ldots x_n) . \tag{2}$$

A definite surface will then correspond to each arbitrary value of the variable parameter $\lambda$, and furthermore, since the geodesic between $A$ and $B$ is intersected by these surfaces, every point along the geodesic can be associated with a quite definite value of $\lambda$.

We shall now consider two neighbouring surfaces, corresponding to the values $\lambda$ and $(\lambda + d\lambda)$, so that the intercept made by these two surfaces on the geodesic is given by

$$ds = w\, d\lambda , \tag{3}$$

where the value of the proportionality factor $w$ follows from the equation

$$ds^2 = g_{hk}\, dx_h\, dx_k \tag{4}$$

[§ 189 eqn. (5) applied to an $n$-dimensional manifold]. Hence we find

$$w^2 = g_{hk} \frac{dx_h}{d\lambda} \frac{dx_k}{d\lambda} . \tag{5}$$

Let $ds'$ be the intercept made by the two surfaces with the values $\lambda$ and $(\lambda + d\lambda)$ on the *varied* line joining the points $A$ and $B$. Then

$$ds' = w'\, d\lambda , \tag{6}$$

whilst

$$ds' - ds = \delta(ds) \tag{7}$$

and

$$w' - w = \delta w . \tag{8}$$

Hence

$$\delta(ds) = \delta w\, d\lambda ; \tag{9}$$

and, since the order of integration and variation in eqn. (1) can be reversed (because the two operations are mutually independent), eqn. (1) may also be written in the form

$$\int_A^B \delta w\, d\lambda = 0 . \tag{10}$$

[1] In view of its definition, the length of a geodesic is by no means necessarily a minimum, but only an extreme value in general.

In order to calculate the variation of $w$, we vary eqn. (5), thus finding

$$(11)\quad 2w\,\delta w = \frac{dx_h}{d\lambda}\frac{dx_k}{d\lambda}\,\delta g_{hk} + g_{hk}\frac{dx_h}{d\lambda}\,\delta\left(\frac{dx_k}{d\lambda}\right) + g_{hk}\frac{dx_k}{d\lambda}\,\delta\left(\frac{dx_h}{d\lambda}\right).$$

We recognize at once that the two last double summands are identical, on account of the symmetry of the fundamental tensor, and either of them may be written in the form

$$g_{hp}\frac{dx_h}{d\lambda}\,\delta\left(\frac{dx_p}{d\lambda}\right).$$

Again, inasmuch as the metric fundamental tensor is a function of the coordinates,

$$(12)\qquad \delta g_{hk} = \frac{\partial g_{hk}}{\partial x_p}\,\delta x_p .$$

Eqn. (11) may therefore be written as follows :—

$$(13)\quad \delta w = \frac{1}{2w}\frac{dx_h}{d\lambda}\frac{dx_k}{d\lambda}\frac{\partial g_{hk}}{\partial x_p}\,\delta x_p + \frac{g_{hp}}{w}\frac{dx_h}{d\lambda}\,\delta\left(\frac{dx_p}{d\lambda}\right).$$

Now

$$(14)\qquad \delta\left(\frac{dx_p}{d\lambda}\right) = \frac{d\,(\delta x_p)}{d\lambda},$$

and hence [2]

$$(15)\quad \frac{g_{hp}}{w}\frac{dx_h}{d\lambda}\,\delta\left(\frac{dx_p}{d\lambda}\right) = \frac{d}{d\lambda}\left(\frac{g_{hp}}{w}\frac{dx_h}{d\lambda}\,\delta x_p\right) - \frac{d}{d\lambda}\left(\frac{g_{hp}}{w}\frac{dx_h}{d\lambda}\right)\delta x_p .$$

If we multiply this equation by $d\lambda$ and integrate between the limits $A$ and $B$, the first term on the right-hand side vanishes, because the variation of the coordinates vanishes both for $A$ and for $B$, seeing that the geodesic and the varied curve coincide at their initial and final points.

Consequently, on substituting the value given by eqn. (15) in eqn. (13), and replacing $\delta w$ in eqn. (10) by the value thus obtained, we find, on reversing the sign,

$$(16)\qquad \int \Theta_p\,d\lambda\,\delta x_p = 0 ,$$

where $\Theta_p$ is given by the relation

$$(17)\qquad \Theta_p = \frac{d}{d\lambda}\left(\frac{g_{hp}}{w}\frac{dx_h}{d\lambda}\right) - \frac{1}{2w}\frac{dx_h}{d\lambda}\frac{dx_k}{d\lambda}\frac{\partial g_{hk}}{\partial x_p}.$$

Inasmuch as the variation of the coordinates is quite arbitrary, however, eqn. (16) can be satisfied only when

$$(18)\qquad \Theta_p = 0 .$$

[2] In accordance with the formula

$$A\,dB = d(AB) - B\,dA.$$

The equations obtained by giving $p$ the values 1, 2, 3, . . . $n$ in an $n$-dimensional geometry represent the equations of the geodesic.

In our previous considerations the parameter has been quite arbitrary. We can therefore stipulate that $\lambda$ shall be equal to the *segments* of the geodesic, as measured from the point $A$. In that case $ds$ is equal simply to $d\lambda$, and hence $w$ is equal to unity. Eqns. (17) and (18) then assume the simpler form

$$g_{hp}\frac{d^2x_h}{ds^2} + \frac{dg_{hp}}{ds}\frac{dx_h}{ds} - \frac{1}{2}\frac{dx_h}{ds}\frac{dx_k}{ds}\frac{\partial g_{hk}}{\partial x_p} = 0\,. \tag{19}$$

But

$$\frac{dg_{hp}}{ds} = \frac{\partial g_{hp}}{\partial x_k}\frac{dx_k}{ds}\,. \tag{20}$$

Furthermore, we can utilize the identity

$$\frac{\partial g_{hp}}{\partial x_k}\frac{dx_k}{ds}\frac{dx_h}{ds} = \frac{\partial g_{kp}}{\partial x_h}\frac{dx_h}{ds}\frac{dx_k}{ds}\,, \tag{21}$$

for the left-hand side of this relation does not change in value if merely other letters are substituted for the indices, and $h$ and $k$ are replaced by $k$ and $h$ respectively. Thus we can equate the second term in eqn. (19) to half the sum of the two expressions forming the identity (21). Eqn. (19) may therefore be written in the form

$$g_{hp}\frac{d^2x_h}{ds^2} + \left[\begin{matrix} h\ k \\ p \end{matrix}\right]\frac{dx_h}{ds}\frac{dx_k}{ds} = 0\,, \tag{22}$$

where the meaning of the expression in square brackets, which is known as *Christoffel's three-indices symbol*, is given by

$$\left[\begin{matrix} h\ k \\ p \end{matrix}\right] = \frac{1}{2}\left(\frac{\partial g_{hp}}{\partial x_k} + \frac{\partial g_{kp}}{\partial x_h} - \frac{\partial g_{hk}}{\partial x_p}\right). \tag{23}$$

On account of the symmetry of the fundamental tensor, it follows at once from this definition that the three-indices symbol is also *symmetrical;* its value is not altered by an interchange of the two upper indices.

In order to simplify subsequent calculations, we shall now introduce another, likewise symmetrical, three-indices symbol, defining it by the relation

$$\left\{\begin{matrix} h\ k \\ r \end{matrix}\right\} = g^{rp}\left[\begin{matrix} h\ k \\ p \end{matrix}\right]. \tag{24}$$

Let us then multiply eqn. (22) by $g^{rp}$ and sum for $p$. According to § 191 eqn. (5), the first term becomes

$$g_h^r \frac{d^2x_h}{ds^2},$$

where, just as in eqn. (22), we have to sum for the index $h$. The summation for $h$, however, simply gives $d^2x_r/ds^2$, because [according to § 191 eqn. (6) ] $g_h^r$ only differs from zero when $h$ and $r$ are identical, and it is then equal to unity. Thus the *equation of a geodesic* finally becomes

$$\frac{d^2x_r}{ds^2} = -\left\{ \begin{matrix} h\,k \\ r \end{matrix} \right\} \frac{dx_h}{ds}\frac{dx_k}{ds}. \tag{25}$$

## § 193. The Riemann Tensor.

Just as the operation of contraction enables us to derive from an arbitrary tensor a tensor of lower rank, so of course, conversely, it must also be possible to *derive* from an arbitrary tensor, by means of purely *mathematical operations*, a *tensor of higher rank*. For this purpose we shall start with a *scalar*, as a tensor of zero rank. Let us assume $\varphi$ to be any scalar function of the coordinates, *i.e.*, to be *invariant* with respect to coordinate transformations. Let us further imagine a *geodesic* constructed quite independently of any coordinate system, and let us denote by $s$ the length of the geodesic (again independently of any coordinate system) as measured from a definite point on it. Then it is evident that the differential coefficient $d\varphi/ds$ must likewise be a quantity independent of the coordinate system. On the other hand, if, as we have assumed, $\varphi$ is a function of the coordinates, we have, of course,

$$\frac{d\varphi}{ds} = \frac{\partial\varphi}{\partial x_h}\frac{dx_h}{ds}. \tag{1}$$

It follows from the definition of a covariant vector that, in accordance with § 190 eqn. (10) (since the $dx_h$ are components of a contravariant vector), the four quantities

$$A_h = \frac{\partial\varphi}{\partial x_h} \tag{2}$$

are the components of a *covariant vector*.[1]

Let us now form the second differential coefficient of $\varphi$

[1] This vector is the gradient of $\phi$.

with respect to $s$; this must, of course, also be *invariant*. Denoting the first derivative by $\psi$, we have

$$\frac{d^2\varphi}{ds^2} = \frac{d\psi}{ds} = \frac{\partial\psi}{\partial x_k}\frac{dx_k}{ds}. \tag{3}$$

But, by eqn. (1),

$$\frac{\partial\psi}{\partial x_k} = \frac{\partial^2\varphi}{\partial x_h\,\partial x_k}\frac{dx_h}{ds} + \frac{\partial\varphi}{\partial x_h}\frac{\partial}{\partial x_k}\left(\frac{dx_h}{ds}\right). \tag{4}$$

Hence

$$\frac{d^2\varphi}{ds^2} = \frac{\partial^2\varphi}{\partial x_h\,\partial x_k}\frac{dx_h}{ds}\frac{dx_k}{ds} + \frac{\partial\varphi}{\partial x_h}\frac{d^2x_h}{ds^2}. \tag{5}$$

Since the right-hand side of this equation contains two entirely independent summands, it is merely a matter of notation to replace the index $h$ in the last term by the index $r$, and this does not involve a similar alteration in the preceding term. Inasmuch, however, as $ds$ represents the element of a geodesic (and the $dx_h$ are the components of this element), the equation of a geodesic [§ 192 eqn. (25)] must be applicable. Consequently, eqn. (5) becomes

$$\frac{d^2\varphi}{ds^2} = \left(\frac{\partial^2\varphi}{\partial x_h\,\partial x_k} - \left\{\begin{matrix} h\,k \\ r \end{matrix}\right\}\frac{\partial\varphi}{\partial x_r}\right)\frac{dx_h}{ds}\frac{dx_k}{ds}. \tag{6}$$

The left-hand side of this equation is invariant. Hence, according to the definitions of tensors, the expression in round brackets must represent the components of a *covariant tensor*:[2] we shall denote these components by $A_{hk}$.[3] Taking eqn. (2) into account, we find

$$A_{hk} = \frac{\partial A_h}{\partial x_k} - \left\{\begin{matrix} h\,k \\ r \end{matrix}\right\} A_r. \tag{7}$$

This tensor is called the *covariant derivative* of the vector $A$.

If, now, we multiply all the components of this tensor by all the components of a covariant vector $B$, we obtain a covariant *tensor of the third rank*. We shall denote its components by

$$C_{hkp} = A_{hp}\,B_k. \tag{8}$$

Hence, according to eqn. (7),

$$C_{hkp} = B_k\frac{\partial A_h}{\partial x_p} - \left\{\begin{matrix} h\,p \\ r \end{matrix}\right\} A_r\,B_k. \tag{9}$$

[2] For the $dx/ds$ are the components of a contravariant vector.

[3] The denotation both of the tensor of the second rank and of the vector by the letter $A$ should not cause any misunderstanding. The distinction between the two is sufficiently expressed by the fact that the tensor has two indices, whereas the vector has only one.

The first term on the right-hand side of this equation suggests that we should define another tensor of the third rank in a similar manner, with the help of the equation

$$D_{hkp} = B_{kp}\, A_h , \tag{10}$$

the tensor $B_{kp}$ being derived from the vector $B_k$ by means of the differential operation expressed by eqn. (7). Thus we find the value of $B_{kp}$ on substituting $B$ for $A$ in eqn. (7), and therefore

$$D_{hkp} = A_h \frac{\partial B_k}{\partial x_p} - \left\{ \begin{matrix} k\ p \\ r \end{matrix} \right\} A_h\, B_r . \tag{11}$$

Next, we define a tensor of the second rank by means of the relation

$$E_{hk} = A_h\, B_k , \tag{12}$$

and a tensor of the third rank with the help of the formula

$$E_{hkp} = C_{hkp} + D_{hkp} . \tag{13}$$

Then we obtain

$$E_{hkp} = \frac{\partial E_{hk}}{\partial x_p} - \left\{ \begin{matrix} h\ p \\ r \end{matrix} \right\} E_{rk} - \left\{ \begin{matrix} k\ p \\ r \end{matrix} \right\} E_{hr} . \tag{14}$$

The tensor of the third rank, $E_{hkp}$, is known as the *covariant derivative* of the tensor of the second rank, $E_{hk}$. Admittedly, eqn. (14) has been derived on the assumption that the components $E_{hk}$ can be represented as the products of the components of two vectors. It follows from the remarks in § 190, however, that any arbitrary tensor of the second rank can be referred to one so represented, and hence the process described by eqn. (14) enables us to derive a tensor of the third rank from any arbitrary tensor of the second rank.

Let us, therefore, replace $E_{hk}$ in eqn. (14) by the tensor $A_{hk}$ as defined in eqn. (7); its covariant derivative may be denoted by $A_{hkp}$. We shall, however, write eqn. (7) in a somewhat different form, replacing the index $r$ by the index $s$; this substitution is necessary to avoid complications in the following derivation. Thus we replace $E_{hk}$ in eqn. (14) by $A_{hk}$, where

$$A_{hk} = \frac{\partial A_h}{\partial x_k} - \left\{ \begin{matrix} h\ k \\ s \end{matrix} \right\} A_s , \tag{15}$$

thereby finding

$$\begin{aligned} A_{hkp} = {} & \frac{\partial^2 A_h}{\partial x_k\, \partial x_p} - \frac{\partial}{\partial x_p} \left\{ \begin{matrix} h\ k \\ s \end{matrix} \right\} A_s - \frac{\partial A_s}{\partial x_p} \left\{ \begin{matrix} h\ k \\ s \end{matrix} \right\} - \left\{ \begin{matrix} h\ p \\ r \end{matrix} \right\} \frac{\partial A_r}{\partial x_k} \\ & + \left\{ \begin{matrix} h\ p \\ r \end{matrix} \right\} \left\{ \begin{matrix} r\ k \\ s \end{matrix} \right\} A_s - \left\{ \begin{matrix} k\ p \\ r \end{matrix} \right\} \frac{\partial A_h}{\partial x_r} + \left\{ \begin{matrix} k\ p \\ r \end{matrix} \right\} \left\{ \begin{matrix} h\ r \\ s \end{matrix} \right\} A_s . \end{aligned} \tag{16}$$

If now, in this equation, we interchange the indices $k$ and $p$, the first term and also, because of the symmetry of the three-indices symbols, the sixth and seventh (last) terms on the right-hand side remain unaltered. The sum of the third and fourth terms is likewise unchanged. Hence, if we suppose eqn. (16) to be formed for $A_{hpk}$, and the resulting equation to be subtracted from eqn. (16), only those terms in which $A_s$ occurs as a factor will remain. Thus we may put

$$A_{hkp} - A_{hpk} = R^s_{hkp} A_s . \tag{17}$$

According to the definitions of tensors, the quantity $R^s_{hkp}$, as is expressed in that symbol itself, must be a *mixed tensor of the fourth rank*, being contravariant as regards the index $s$, and covariant as regards the other indices.

From eqn. (16) we find

$$R^s_{hkp} = \frac{\partial}{\partial x_k}\left\{\begin{matrix} h\,p \\ s \end{matrix}\right\} - \frac{\partial}{\partial x_p}\left\{\begin{matrix} h\,k \\ s \end{matrix}\right\} + \left\{\begin{matrix} h\,p \\ r \end{matrix}\right\}\left\{\begin{matrix} r\,k \\ s \end{matrix}\right\} - \left\{\begin{matrix} h\,k \\ r \end{matrix}\right\}\left\{\begin{matrix} r\,p \\ s \end{matrix}\right\}. \tag{18}$$

It follows from the meaning of the Christoffel's three-indices symbols [§ 192 eqns. (23) and (24) ] that the tensor of the fourth rank is completely determined as soon as we know the functional relations connecting the components of the metric fundamental tensor and the coordinates. If the metric field is constant, so that the fundamental tensor is independent of the coordinates, the tensor of the fourth rank vanishes; and, indeed, since it is independent of the coordinate system, it vanishes for *all* coordinate systems. In this lies the great importance of the mixed tensor of the fourth rank, which was first derived by *Riemann*, and is therefore known as the *Riemann tensor*. The *vanishing* of the Riemann tensor is thus the *necessary condition for the constancy of the metric field*.[4]

Like any other mixed tensor, the Riemann tensor can also, of course, be contracted. On carrying out this operation, by putting the index $s$ equal to $p$ and summing for $p$, we obtain as the formula for the *contracted Riemann tensor*

$$R_{hk} = \frac{\partial}{\partial x_k}\left\{\begin{matrix} h\,p \\ p \end{matrix}\right\} - \frac{\partial}{\partial x_p}\left\{\begin{matrix} h\,k \\ p \end{matrix}\right\} + \left\{\begin{matrix} h\,p \\ r \end{matrix}\right\}\left\{\begin{matrix} r\,k \\ p \end{matrix}\right\} - \left\{\begin{matrix} h\,k \\ r \end{matrix}\right\}\left\{\begin{matrix} r\,p \\ p \end{matrix}\right\}. \tag{19}$$

Furthermore, according to § 190 eqn. (13), we can derive from the contracted Riemann tensor a scalar invariant, by means of the relation

$$R = g^{hk} R_{hk} . \tag{20}$$

[4] It may be shown that this condition is also sufficient.

The quantity $R$ is known as the *Riemann invariant.* Finally, we can obtain the *contravariant components* corresponding to the covariant components of the contracted Riemann tensor, with the aid of the equation [according to § 191 eqn. (11) ]

$$R^{ps} = g^{pk}\, g^{sh}\, R_{hk}\,. \tag{21}$$

## § 194. Einstein's and Newton's Laws of Gravitation.

If the Riemann invariant be integrated over an arbitrarily delimited region of the geometrical manifold, and the *variation of this integral* be then formed in such a way that the fundamental tensor is varied only within the boundaries of this region, purely mathematical considerations (which are too complicated to be reproduced here) lead to the components of a tensor

$$S_{hk} = R_{hk} - \frac{1}{2} g_{hk}\, R\,. \tag{1}$$

It is this "*curvature tensor*" to which Einstein's theory *equates* the *matter tensor*, apart from a universal proportionality factor.[1]

In accordance with § 187, the components of the matter tensor are found by forming the products of the quantities $dx_h/ds$ in pairs, and then multiplying by the stationary density ($\varrho_0$) of the proper mass. Since the quantities $dx_h/ds$ are contravariant [according to § 190 eqns. (6) and (9) ], whereas $\varrho_0$ is a scalar, we obtain in this way the *contravariant components of the matter tensor*

$$T^{hk} = \varrho_0 \frac{dx_h}{ds} \frac{dx_k}{ds}\,. \tag{2}$$

In accordance with Einstein's theory, the relation

$$T_{hk} = S_{hk} \tag{3}$$

must be satisfied ; *i.e.*,

$$T_{hk} = R_{hk} - \frac{1}{2} g_{hk}\, R\,, \tag{4}$$

or,

$$\varrho_0 \frac{dx_h}{ds} \frac{dx_k}{ds} = R^{hk} - \frac{1}{2} g^{hk}\, R\,, \tag{5}$$

[1] If the variation vanishes at the boundary, it may be shown that

$$\delta \int R\, d\omega = \int \left( R_{hk} - \tfrac{1}{2} g_{hk}\, R \right) \delta g^{hk}\, d\omega\,.$$

the contravariant components of the contracted Riemann tensor being determined by § 193 eqn. (21).

Eqn. (4) leads directly to a relation which is important for subsequent considerations. It follows from § 190 eqn. (13) that a scalar invariant, which may be denoted by $T$, can be formed by means of the relation

$$T = g^{hk} T_{hk} . \tag{6}$$

Again [by § 193 eqn. (20)], $g^{hk} R_{hk}$ is equal to $R$, and [by § 191 eqn. (2)] $g^{hk} g_{hk}$ is equal to four. Hence, on multiplying eqn. (4) by $g^{hk}$ and summing for $h$ and $k$, we find

$$T = R - 2R ,$$

or,

$$R = - T . \tag{7}$$

Einstein's Law of Gravitation postulates that both the "*field equations*" (5) and the equation which expresses the *geodesic character of the world-line representing the motion* shall be simultaneously satisfied. Thus, by § 192 eqn. (25), the relation

$$\frac{d^2 x_r}{ds^2} = - \left\{ \begin{matrix} h\ k \\ r \end{matrix} \right\} \frac{dx_h}{ds} \frac{dx_k}{ds} \tag{8}$$

must also exist. The two eqns. (5) and (8), which must be simultaneously satisfied, therefore form the analytical expressions of Einstein's Law of Gravitation.

In discussing these equations, we shall first consider a special case in which three greatly simplifying *assumptions* are satisfied. The *first* assumption is that the *deviations from Euclidean geometry* are *only small*. Hence, using a rectangular coordinate system, the quantities $g_{hk}$ shall differ but little from 1 or 0, according as $k$ is or is not equal to $h$. According to the definitions of cofactors and the contravariant fundamental tensor, the quantities $g^{hk}$ also will then differ but little from 1 or 0, again according as $k$ is or is not equal to $h$.[2] Our first assumption may therefore be expressed by the formula

$$\left. \begin{matrix} g_{hk} \\ g^{hk} \end{matrix} \right\} \backsimeq \delta^k_h , \tag{9}$$

where $\delta^k_h$ denotes the mixed fundamental tensor, the components of which are always unity or zero [§ 191 eqn. (6)].

If, now, eqn. (9) be satisfied, the three-indices symbols

[2] This follows also from scheme (10) in § 189.

in curly brackets can also be reduced to a simpler form. In accordance with § 192 eqn. (24), which defines this symbol,

$$\left\{ \begin{matrix} h\,k \\ r \end{matrix} \right\} = g^{rp} \left[ \begin{matrix} h\,k \\ p \end{matrix} \right].$$

Of the sum of the four terms on the right-hand side, three summands, however, can [by eqn. (9)] be neglected. We have only to consider the summand in which $p$ is equal to $r$, and then, by eqn. (9), $g^{pr}$ is simply equal to unity, so that, by way of approximation

$$\left\{ \begin{matrix} h\,k \\ r \end{matrix} \right\} = \left[ \begin{matrix} h\,k \\ r \end{matrix} \right]. \tag{10}$$

In the special case considered, therefore, the curly-bracketed three-indices symbols can be replaced simply by the square-bracketed symbols.

Since the three-indices symbols contain only the partial derivatives of the $g_{hk}$ with respect to the coordinates, it follows further from eqn. (9) that, as a result of our assumption, the *three-indices symbols* must be *small quantities*, small, at least, in the first order.

Our *second assumption* is to the effect that the *mechanical velocities* involved are *small as compared with the velocity of light*. In that case it follows from § 178 eqns. (2) and (5) that

$$\frac{dx_1}{ds}, \frac{dx_2}{ds}, \frac{dx_3}{ds} \text{ are small}; \quad \frac{dx_4}{ds} \backsimeq 1. \tag{11}$$

Moreover, according to § 187 eqn. (3),

$$T = \varrho, \tag{12}$$

and finally (inasmuch as $dx_4/ds$ is approximately equal to unity)

$$\frac{d^2x_r}{ds^2} = -\frac{1}{c^2}\frac{d^2x_r}{dt^2}. \tag{13}$$

Finally, the *third* simplifying assumption we shall make is that the *gravitational field* shall be *static*, *i.e.*, the components $g_{hk}$ shall be *spatially*, but *not temporally*, *variable*. The partial derivatives with respect to $x_4$ may thus be simply neglected, and we put

$$\frac{\partial g_{hk}}{\partial x_4} = 0. \tag{14}$$

When our three assumptions hold good, in the first place

the equation of a geodesic takes a very simple form. The expression

$$\left\{ \begin{matrix} h\,k \\ r \end{matrix} \right\} \frac{dx_h}{ds} \frac{dx_k}{ds}, \tag{15}$$

it is true, represents a double sum of 16 summands (since we have to sum both for $h$ and for $k$). According to eqn.(11), however, 15 of these summands can be neglected in comparison with the one term

$$\left\{ \begin{matrix} 4\,4 \\ r \end{matrix} \right\} \frac{dx_4}{ds} \frac{dx_4}{ds}.$$

Now it follows from eqn. (10) and the definition of the three-indices symbols [§ 192 eqn. (23) ] that,

$$\left\{ \begin{matrix} 4\,4 \\ r \end{matrix} \right\} = \left[ \begin{matrix} 4\,4 \\ r \end{matrix} \right] = \frac{1}{2} \left( \frac{\partial g_{4r}}{\partial x_4} + \frac{\partial g_{r4}}{\partial x_4} - \frac{\partial g_{44}}{\partial x_r} \right),$$

or since, by eqn. (14), the partial derivatives with respect to $x_4$ vanish,

$$\left\{ \begin{matrix} 4\,4 \\ r \end{matrix} \right\} = - \frac{1}{2} \frac{\partial g_{44}}{\partial x_r}. \tag{16}$$

If, further, we take into account eqns. (11) and (13), the equation of a geodesic becomes, in the special case considered, simply

$$\frac{d^2 x_r}{dt^2} = - \frac{c^2}{2} \frac{\partial g_{44}}{\partial x_r}, \tag{17}$$

and, in conjunction with this equation, we have still to take eqn. (5), which, on the basis of the same assumptions, is likewise considerably simplified.

In the first place, the expression for the contracted Riemann tensor can be greatly simplified. This expression [§ 193 eqn. (19) ] consists of four terms. The last two of these, which represent the products of two three-indices symbols, may be omitted, because, in accordance with our previous remarks, these symbols are small quantities, at least in the first order. Furthermore, if we put the index $k$ equal to 4, then, by eqn. (14), the first term may also be neglected, and only the second remains. We have therefore

$$R_{h4} = - \frac{\partial}{\partial x_p} \left\{ \begin{matrix} h\,4 \\ p \end{matrix} \right\}. \tag{18}$$

Hence, by eqn. (16),

$$R_{44} = \frac{1}{2} \left( \frac{\partial^2 g_{44}}{\partial x_1^{\,2}} + \frac{\partial^2 g_{44}}{\partial x_2^{\,2}} + \frac{\partial^2 g_{44}}{\partial x_3^{\,2}} + \frac{\partial^2 g_{44}}{\partial x_4^{\,2}} \right) \tag{19}$$

According to eqn. (14), the partial differential coefficient with respect to $x_4$ vanishes. Moreover, the sum of the second partial derivatives with respect to the spatial coordinates represents the well-known Laplace derivative, generally denoted by the symbol $\nabla^2$ (§ 32). Hence

$$R_{44} = \tfrac{1}{2}\nabla^2 g_{44}\,. \tag{20}$$

We shall next consider the contravariant components of the contracted Riemann tensor ; according to § 191 eqn. (11),

$$R^{44} = g^{4h}\,g^{4k}\,R_{hk}\,, \tag{21}$$

It follows, however, from eqn. (9) that, in this double sum of 16 summands, all other summands can be neglected in comparison with the one for which $h = k = 4$. Hence, since $g_{44}$ is approximately equal to unity,

$$R^{44} = R_{44}\,. \tag{22}$$

Moreover, by eqns. (7) and (12),

$$R = -\varrho\,. \tag{23}$$

Let us now put $h = k = 4$ in eqn. (5). According to eqn. (11), the differential coefficients then become equal to unity, whilst $\varrho_0$ is equal to $\varrho$, and thus we obtain the relation

$$\varrho = \tfrac{1}{2}\nabla^2 g_{44} + \tfrac{1}{2}\varrho\,,$$

or,

$$\varrho = \nabla^2 g_{44}\,. \tag{24}$$

This equation represents a partial differential equation of the second order, and it follows from § 51 eqn. (10) that its solution is given by the expression

$$g_{44} = -\frac{1}{4\pi}\int\frac{\varrho\,d\tau}{r}\,, \tag{25}$$

where $r$ denotes the distance of the volume-element $d\tau$ from the field-point, and $\varrho$ the mass density in the volume-element. Finally, if we combine eqn. (25) with eqn. (17) (in which we shall replace the index $r$ by the index $h$),[3] we obtain the *equation of motion in a material field* in the form

$$\frac{d^2x_h}{dt^2} = \frac{c^2}{8\pi}\frac{\partial}{\partial x_h}\int\varrho\,\frac{d\tau}{r}\,. \tag{26}$$

Hitherto we have measured mass in such a way that the universal proportionality factor between the curvature tensor and the matter tensor is simply equal to unity. When, however, we measure mass density, as is usual, in gm./cm.$^3$,

[3] In order that the letter $r$ shall not occur with two different meanings in the same equation.

$\varrho$ will be replaced in the above formulæ by the product $f\varrho$, where $f$ denotes a universal proportionality factor. We shall introduce a new universal constant, defined by the relation

$$\kappa = \frac{f c^2}{8\pi}; \tag{27}$$

then

$$\frac{d^2 x_h}{dt^2} = \kappa \frac{\partial}{\partial x_h} \int \frac{\varrho \, d\tau}{r}. \tag{28}$$

We shall now suppose the distribution of mass, previously considered continuous, to be only approximately so; we imagine in reality a number of discrete particles to be present, and we shall describe the individually differing masses by $m$, and by $r$ the individually differing distances from the field-point considered. Then we have

$$\int \frac{\varrho \, d\tau}{r} = \sum \frac{m}{r}. \tag{29}$$

Finally, we suppose a particle of mass $m'$ to be situated at the field-point itself. On multiplying eqn. (28) by $m'$, we obtain the relation

$$m' \frac{d^2 x_h}{dt^2} = \frac{\partial}{\partial x_h} \sum \frac{\kappa \, m \, m'}{r}. \tag{30}$$

Now the three quantities $d^2 x_h / dt^2$ are nothing else than the components of the spatial vector $d\mathbf{v}/dt$, where $\mathbf{v}$ is the velocity of the particle $m'$. On the other hand, in accordance with a relation derived elsewhere and frequently used, the three quantities $\partial(1/r)/\partial x_h$ are the components of the vector $(-\mathbf{r}/r^3)$, where $\mathbf{r}$ is the radius-vector drawn from a mass $m$ to the field-point.[4] Eqn. (30) may therefore be written in the form

$$m' \frac{d\mathbf{v}}{dt} = -\sum \frac{\kappa \, m \, m'}{r^2} \cdot \frac{\mathbf{r}}{r}. \tag{31}$$

The left-hand side of this equation may be regarded as representing a *force* exerted on the particle $m'$ in consequence of the presence of the other particles. Accordingly, the right-hand side may be interpreted as the resultant of the individual forces acting between the particle $m'$ and all the other

[4] For the $x_h$ are the coordinates of the field-point. Hence § 50 eqn. (4) applies, according to which, for a variable field-point,

$$\operatorname{grad}_a \left(\frac{1}{r}\right) = -\frac{1}{r^2}\frac{\mathbf{r}}{r},$$

where $\mathbf{r}$ denotes the radius-vector drawn from the source to the field-point.

particles. As the equation shows us, these forces must act *in the direction of the line joining the particles* concerned, and, inasmuch as they are opposite in sign to the radius-vector drawn from $m$ to $m'$, they will tend to draw $m'$ towards $m$, *i.e.*, they must be *forces of attraction*. The magnitude of the force is, apart from a universal proportionality factor, *equal to the product of the two masses divided by the square of the distance between them.*

Thus, in consequence of the presence of the other particles, every particle moves as if a *universal force of attraction* were at work *between all masses*, governed by the law just given. This, however, is none other than the well-known *Law of Gravitation*, established by *Newton*.

## § 195. The Metrics of Space and Time in a Gravitational Field.

The derivation of Newton's from Einstein's Law of Gravitation, as a first approximation, in itself constitutes a great success for Einstein's Theory. The possibility of experimentally testing this new theory, however, is to be sought in the appearance of several phenomena which, whilst necessarily following from the more complete Law of Einstein, cannot be explained on the basis of Newton's Law.

Einstein's gravitational equations are partial *differential equations* of the second order, in which the components of the metric fundamental tensor are differentiated with respect to the coordinates. The general problem would then consist in calculating the quantities $g_{ik}$ as functions of the world-coordinates, for a given space-time distribution of matter. The rigid solution of this problem is so complicated that it has not yet been accomplished. *Einstein*, and in a more accurate manner, *Schwarzschild*, however, have obtained the solution for the special case in which the *gravitational field is produced by a single particle at rest*, when the field must obviously possess spherical symmetry in space.

The calculations, which are too involved to be reproduced here, lead to the following solution: If we denote by $M$ the single mass producing the field, and by $\kappa$ the universal gravitational constant, and if we put, for brevity,

$$\alpha = \frac{\kappa M}{c^2}, \tag{1}$$

then, at a matter-free point with the spatial coordinates $x_1$,

$x_2$, $x_3$, at a distance $r$ from the particle generating the field, we find

$$(2)\qquad \left\{\begin{aligned} g_{hk} &= \delta^k_h + \frac{2\alpha x_h x_k}{r^3}, \\ g_{h4} &= g_{4h} = 0, \\ g_{44} &= 1 - \frac{2\alpha}{r}. \end{aligned}\right\} \qquad \begin{matrix}(h = 1, 2, 3)\\(k = 1, 2, 3)\end{matrix}$$

$$\delta^h_k = \begin{cases} 1 & (h = k) \\ 0 & (h \neq k) \end{cases}$$

Since the dimensions of the gravitational constant are gm.$^{-1}$ cm.$^3$ sec.$^{-2}$, the dimensions of the quantity $\alpha$ defined by eqn. (1) are those of a length, and $\alpha$ is therefore called the *gravitational radius* of the mass. Now

$$(3)\qquad \kappa = 6{\cdot}68 \times 10^{-8}\ \text{gm.}^{-1}\ \text{cm.}^3\ \text{sec.}^{-2},$$

and thus we obtain the gravitational radius of a mass in centimetres on multiplying the number of grams by $7 \times 10^{-29}$. For example, the gravitational radius of an electron is about $6 \times 10^{-56}$ cm.—an inconceivably small quantity. The earth's mass is about $6 \times 10^{27}$ gm., so that its gravitational radius is approximately 4 mm. The sun is about 330,000 times as heavy as the earth, and thus its gravitational radius would be rather more than 1 km.

Eqns. (2) only hold good in space outside the mass producing the field. As eqns. (2) show us, the deviations from Euclidean metrics are thus determined by the *ratio* between the gravitational radius and the radius of the mass (assumed to be spherical). For an electron of about $10^{-13}$ cm. radius the ratio would be unimaginably small, *viz.*, about $10^{-43}$ in order of magnitude. For the earth the ratio would be about $6 \times 10^{-10}$, and for the sun approximately $2 \times 10^{-6}$. An appreciable influence of a gravitational field on the metrics of space and time is thus scarcely to be expected in atomic processes ; it is, in fact, hardly to be expected on the earth at all, but rather on the sun and in its immediate vicinity, as well as on other large stars. An experimental confirmation of Einstein's gravitational theory is thus to be sought, not in laboratory researches, but in astronomical measurements.

For the purposes of a closer investigation of the metrics in a gravitational field, let us consider quite generally two adjacent events in time and space, the corresponding world-coordinates being $x$, $y$, $z$, $i\,c\,t$ and $(x + dx)$, $(y + dy)$, $(z + dz)$, $i\,c\,(t + dt)$ respectively. The four-dimensional interval $(ds)$ between the two world-points is, as we know,

invariant and hence independent of the gravitational field. According to the generalized Theorem of Pythagoras [§ 189 eqn. (5)] we have, since by eqns. (2) the $g_{h4}$ vanish,

$$(4) \qquad ds^2 = \sum_{h=1}^{h=3} \sum_{k=1}^{k=3} g_{hk}\, dx_h\, dx_k + g_{44}\, dx_4{}^2,$$

where

$$dx_4 = i\, c\, dt.$$

When, in particular, the two events are simultaneous, then $ds$ represents the spatial interval between the two points at which the two events occur. We may distinguish between two special cases. In the first case let both points lie on the $x$-axis at distances $r$ and $(r + dr)$ respectively from the particle producing the field. Then

$$(5) \qquad x_1 = r, \quad x_2 = x_3 = 0, \quad dx_2 = dx_3 = dt = 0,$$

so that we have simply

$$(6) \qquad ds^2 = g_{11}\, dx_1{}^2,$$

or,

$$dx_1 = \frac{ds}{\sqrt{g_{11}}},$$

or finally, by way of approximation and according to eqns. (2),

$$(7) \qquad dx_1 = \frac{ds}{\sqrt{1 + \dfrac{2\alpha}{r}}} = ds\left(1 - \frac{\alpha}{r}\right).$$

In this case, therefore, a *measuring-rod* lying in the $x$-direction, or *radially*, is apparently *shortened by the gravitational field.*

In the second special case we shall suppose the rod to be lying *tangentially* at the same point, but parallel, say, to the $y$-axis. Then

$$(8) \qquad x_1 = r, \quad x_2 = x_3 = 0, \quad dx_1 = dx_3 = dt = 0,$$

so that

$$(9) \qquad ds^2 = g_{22}\, dx_2{}^2.$$

But, by eqns. (2) and (8),

$$(10) \qquad g_{22} = 1 + \frac{2\alpha x_2{}^2}{r^3} = 1.$$

Hence the *length* of a rod lying *tangentially* in the gravitational field is *not affected* at all. A metre-rule must

experience a contraction of about 6 Å. when raised from a horizontal into a vertical position on the earth, but this contraction is only about one-tenth of the Lorentz contraction due to the earth's motion.

When the two events represented by the two previously considered adjacent world-points occur at the same place, $ds$ denotes the time-interval between the two events, multiplied by $ic$. Thus we have

$$ds^2 = g_{44}\, dx_4^2 , \tag{11}$$

or, by eqns. (2),

$$dx_4 = \frac{ds}{\sqrt{1 - \frac{2\alpha}{r}}} = ds\left(1 + \frac{\alpha}{r}\right). \tag{12}$$

Hence, since $dx_4$ is larger than $ds$, *all processes* are apparently *retarded in a gravitational field.*

The alteration of the metrics of space and time in a gravitational field must obviously also have the effect that, *in a gravitational field, rays of light* are not straight but *curved,* as we might surmise without entering upon any calculations. Accurate calculations (too long to be given here) lead, in fact, to the following result :—When a ray of light passes by a gravitational mass of gravitational radius $\alpha$ at a distance $r$ from it (measured from the centre of the mass), it experiences a *deflection,* of which the total magnitude [1] in circular measure is equal to $4\alpha/r$.

When the alterations of the metrics in a gravitational field are taken into account, *Kepler's Laws* also assume a *more complicated form.* Exact calculations (too involved to be reproduced here) show that the major axis of the planetary orbits must slowly rotate in the orbital plane, in the sense in which the planet revolves. The *perihelion of the planet's orbit* must therefore *advance* slowly, and the angle of advance for each revolution is found by accurate calculations to be given in circular measure by

$$\varphi = \frac{6\pi\alpha}{a(1-\epsilon^2)}; \tag{13}$$

in this expression $\alpha$ is the gravitational radius of the sun, $a$ the semi-major axis, and $\epsilon$ the eccentricity of the orbit.

[1] As Einstein discovered in 1911, half of this amount is due to the fact that the ray of light is of itself deflected in the gravitational field on account of its energy ; this follows from the two laws of the inertia of energy and of the identity of inertial and gravitational mass.

## § 196. The Experimental Confirmation of Einstein's Theory of Gravitation.

The *perihelion motion of the planets* demanded by Einstein's theory, as opposed to the Newtonian theory, was actually discovered in 1845 by *Leverrier*, the astronomer. Leverrier observed this motion in the case of the planet nearest the sun, *Mercury*, but astronomers were unable to give a satisfactory explanation of the phenomenon, which was contradictory to Newton's theory. The perihelion motion is in the sense of the planet's motion, and, as Leverrier observed, amounts to 43 seconds of arc in a century.[1] Exactly the same value, however, was obtained by Einstein from the equation in his theory [§ 195 eqn. (13)], on inserting the known values for the major axis and the eccentricity of the orbit of Mercury, and the gravitational radius of the sun.



FIG. 135.

Einstein's gravitational theory has received a second striking confirmation based on the previously mentioned fact that a ray of light is deflected in passing a gravitational mass; as already stated, the total deflection amounts to $4\alpha/r$. Since, as we have already mentioned, the ratio $\alpha/r$ is equal to $2 \times 10^{-6}$ at the surface of the sun, a *ray of light* which *grazes the sun* must experience a total deflection of $8 \times 10^{-6}$ or (in the usual angular measure) of 1·7 seconds of arc.[2] If the ray passes the sun at a distance equal to $\gamma$ times the sun's radius from its centre, the deflection should be $1{\cdot}7/\gamma''$. In consequence of the deflection of the ray, the star emitting it would not appear in its actual position to a terrestrial observer. Stars which appear close to the sun in the sky must therefore be *apparently displaced* by a very small amount *away from the sun*, the displacement being larger the nearer in line they are to the sun (Fig. 135).

Now the stars nearly in line with the sun are admittedly invisible as a rule, but, during a *total solar eclipse*, they can be observed and photographed. It has, in fact, been possible already to detect the so-called "Einstein effect" in several solar eclipses, the first observations having been made in

[1] This amount represents the excess of the perihelion motion actually observed over that calculated after allowing for all possible perturbations.

[2] Since 360° is equal to $2\pi$ in circular measure, a radian is equal to $360 \times 60 \times 60/2\pi''$.

1919.[3] Exact measurements of the photographs obtained of the sun and the stars surrounding it during the eclipse have confirmed the effect predicted by Einstein, not only qualitatively, but also quantitatively, with complete accuracy within small limits of error.

Finally, Einstein has deduced from the alteration of the metrics of time in a gravitational field that the *frequency of spectral lines* originating in the *stars* must apparently be *lowered* as compared with lines of terrestrial origin—in other words, that a *displacement towards the red* must occur. In point of fact, this prediction of Einstein's theory has been *confirmed*, particularly by the observations (already mentioned in § 91) on the *companion of Sirius*. It should not, however, be overlooked that the *corpuscular theory of light* accounts for this displacement towards the red, and to just the same extent, as we have already seen when discussing that theory (§ 91).

## § 197. Size and Mass of the Universe.

It has long been recognized by physicists and astronomers that, in the classical theory of gravitation, Newton's Law of Gravitation *cannot be reconciled* with the concept of a *spatial infinity of the universe*. This incompatibility appears clearly from a simple consideration. Let us imagine a space the linear dimensions of which are large compared with the distance between two neighbouring fixed stars, but small compared with the extent, say, of the Milky Way. We can regard the quotient of the total mass of all the stars contained in this volume and the volume itself as the *density* in that portion of the universe. We shall now suppose two spherical surfaces to be constructed about any arbitrary point as centre, the radii of these spheres being $r$ and $(r + \Delta r)$, where $r$ is very large and even $\Delta r$ is large as compared with the mean distance between two neighbouring stars. The two spherical surfaces will then form the boundaries of a " spherical shell " (admittedly of very large thickness). If the density in the universe be fairly uniform, inasmuch as no one point in the infinite space appears privileged as regards any other, then the total mass contained within the spherical shell will be

[3] Half of the effect had already been predicted by Einstein in 1911, before the evolution of the general theory of relativity (*cf.* § 195 footnote 1). In August, 1914, it was arranged to send a German expedition to look for the predicted effect in a total eclipse observable in Southern Russia, but the outbreak of war prevented this.

proportional to $r^2 \Delta r$, and the force of attraction exerted at the centre by each mass will be *inversely* proportional to $r^2$. The total attraction exerted by a spherical shell is thus approximately proportional to the thickness of the shell, but independent of $r$ itself. Hence, if we regard space as infinite and containing a uniform distribution of stars, a force of attraction will be exerted at every point from all directions, this force being infinite or at least indefinite, so that a stable solar system would be inconceivable.

On the other hand, if we suppose the universe to be spatially infinite, but the number of stars finite, then, on the ground of the results of the kinetic theory of gases, there could be no reason why the stars (which are comparable to molecules in that theory) should not lose themselves in the infinite space, or why there should ever be a dynamic equilibrium between them.

These difficulties of the classical theory have been surmounted by Einstein, on the assumption that, although the *universe* is *boundless*, yet it has a *finite volume*. The possibility of such an assumption follows directly from *Riemannian geometry*. Just as a circular line has finite length, but no beginning nor end, and just as the earth's surface has no beginning nor end, but yet contains a definite number of square metres, so could a *curved space* (of which we can certainly form no conception) be of finite volume and yet be boundless. A geodesic drawn as a great circle on the earth's surface leads back to the starting-point, and similar considerations would apply to a three-dimensional spherical super-surface, or a so-called *spherical space*. As in the cases of a circle and a spherical surface, we have to assume a *radius of curvature* for a spherical space, and, denoting it by $R$, the relation (which we shall state here without giving the proof) between it and the finite volume is given by the simple expression

$$V = 2\pi^2 R^3 . \tag{1}$$

In consequence of the Principle of Relativity, the idea of space has admittedly lost its independence. On the basis of the important experimental fact that the *relative velocities of stars* are *always small as compared with the velocity of light*, however, we can, to a certain degree of approximation, split up the Minkowski-world into space and time. We can thus speak approximately of a *space*, and this Einstein regards as *spherical* in the sense already mentioned, in that it is boundless, but has a finite volume.

Again, the assumption of a spherically curved space involves the addition of a *supplementary term* in Einstein's Law of Gravitation, but this is of no importance for the derivation of Newton's Law as a special case, or of the three experimentally confirmed effects of the general theory of relativity. The *cosmological law of gravitation* equates the matter tensor, not to the curvature tensor itself, but to the difference between the curvature tensor and the product of the fundamental tensor and a universal constant; in this form also the Law of Gravitation is, of course, independent of the coordinate system. Denoting the universal, so-called *cosmological*, constant by $\lambda$, Einstein's gravitational equation in its covariant form becomes

$$R_{hk} - \frac{1}{2} g_{hk} R - \lambda g_{hk} = T_{hk} . \tag{2}$$

This equation is, of course, supplemented by the one expressing the geodetic nature of a world-line, which is not altered.

If, now, we combine these two equations with the third one, which expresses the existence of spherical metrics (in the sense previously discussed) in space split off from the Minkowski-world, calculation leads to the two relations

$$\lambda = \frac{4\pi\kappa\varrho}{c^2} = \frac{1}{R^2} , \tag{3}$$

where $\kappa$ denotes the gravitational constant, $\varrho$ the *mean mass density in the universe*, and $R$ *the radius of curvature of the universe* [not, as in eqn. (2), the Riemann invariant].

On the other hand, denoting the *total mass of the universe* by $M$, we find by eqn. (1)

$$\varrho = \frac{M}{2\pi^2 R^3} , \tag{4}$$

and hence, by eqn. (3),

$$\frac{4\pi\kappa M}{2\pi^2 c^2 R^3} = \frac{1}{R^2} ,$$

or,

$$M = \frac{\pi c^2 R}{2\kappa} . \tag{5}$$

Since the gravitational radius of the universe is equal to $\kappa M / c^2$, the *radius of curvature of the universe* is *identical with its gravitational radius* divided by the numerical factor $\pi/2$.

On the basis of astronomical estimates, it has been possible to derive the order of magnitude of the *mean density of matter* in the universe, and in this way definite values have also been calculated for the *extent of the universe* (*viz.*, the

largest possible distance within it) and for its mass. *De Sitter*,[1] admittedly after some modification of the bases of Einstein's cosmological theory of gravitation,[2] found a value of 100 million light-years for the *extent of the universe*.[3] A value some thousand times as large was obtained by *Hubble*[4] from investigations of *spiral nebulæ*.[5] By way of comparison, it may be mentioned that the nearest spiral nebula lies at a distance of 850,000 light-years from us, whilst, according to Hubble, the distance of the furthest spiral nebula which is to be observed in the strongest telescopes would be some 150 million light-years.

The *mass of the universe* is given by de Sitter as $10^{54}$, and by Hubble as $10^{57}$ grams. Hubble estimates that the number of spiral nebulæ visible in the large telescope of the Mount Wilson observatory is about two million, with a total mass of, perhaps, $10^{50}$ grams.[6] It follows from this estimate that the *total number of all material primordial particles in the universe* (protons and electrons)[7] would be of the order of magnitude of $10^{81}$.

[1] W. de Sitter, *Roy. Astron. Soc., Monthly Notices*, **78** (1917), p. 229.

[2] Einstein's theory indicates that the Minkowski-world is cylindrical, whereas de Sitter's theory regards it as being spherical.

[3] A light-year is the distance traversed by light in one year: it is equal to $1{\cdot}0 \times 10^{18}$ cm.

[4] E. Hubble, *Astrophys. Journ.*, **64** (1926), p. 369.

[5] Since, in consequence of the curvature of space, every ray of light must return to its source after about 100 million years (de Sitter) or 100,000 million years (Hubble), an object in the northern celestial hemisphere might conceivably be identical with one in the southern: just as a place 10,000 km. distant to the east of any point on the equator would be identical with a place 30,000 km. to the west of that point.

[6] According to recent astronomical estimations, the Milky Way system, which in all probability represents a spiral nebula, contains some 30,000 million stars. Multiplying this number by the mass of the sun ($2 \times 10^{33}$ grams) and the number of visible spiral nebulæ ($2 \times 10^{6}$), we obtain the quoted figure of $10^{50}$ grams. For the rest, it may be mentioned that calculations by Jeans have shown that, of the 30,000 million stars in our stellar system, it is unlikely that more than 10,000 possess a planetary system. The formation of such a system can occur only about once in 1,000 million years in the system of the Milky Way.

[7] The number of protons corresponding to 1 gram is $6 \times 10^{23}$, and, since matter is neutral, the number of electrons is the same.

# APPENDIX

## SUMMARY OF THE CONTENTS OF VOLUME II

### Part III

### ATOMIC THEORY

#### Chapter XI

**Electrons, Atoms, and Light-Quanta**

§ 86. It has been found from the movements executed by small particles, with weak electrical charges, under the dual influence of their own weight and an electric field directed vertically upwards, that the charges on such particles are exact integral multiples of an elementary quantum of electricity ($e$). The value of this quantum is found by observation to be $4{\cdot}77 \times 10^{-10}$ E.S.U.

§ 87. The number by which the atomic weight numbers must be divided in order to obtain the absolute masses of the atoms is termed the Loschmidt number ($L$). The product $L\ e$ is known from electrolytic measurements, on the basis of Faraday's fundamental law, and the value of $L$ thus follows as $6{\cdot}06 \times 10^{23}$. Hence the mass of an atom of hydrogen is $1{\cdot}66 \times 10^{-24}$ gm.

§ 88. The specific charge of cathode ray particles is found from their electric and magnetic deflections to be about 1,800 times that of an ionized atom of hydrogen. On the assumption that the charge on each particle is one elementary quantum, the cathode rays thus consisting of electrons, the mass ($m$) of an electron is found to be $9{\cdot}0 \times 10^{-28}$ gm. The $\beta$-rays emitted by radioactive substances likewise consist of electrons, and their velocities amount to from 30 to 99·8% of the velocity of light. It has been possible to demonstrate experimentally with these rays the increase in mass with the velocity demanded by the theory of relativity.

§ 89. Canal rays and anode rays consist of positive ions, the masses of which are found by positive ray analysis to be always approximately integral multiples of the mass of a hydrogen

atom. The positive $\alpha$-rays emitted by radioactive substances have initial velocities of 5—7% of the velocity of light. The scintillation method allows of a count of the $\alpha$-particles expelled by a preparation in a given time. In this way the charge on a single $\alpha$-particle can be derived, and hence also its mass, since the specific charge is known from deflection experiments. It has been found that the mass of an $\alpha$-particle is four times as large as that of a hydrogen atom, whilst its positive charge amounts to two elementary quanta. In hydrogen, $\alpha$-particles produce long-range H-rays.

§ 90. Careful investigations on the photo-electric effect show that the equivalent potential of the liberated electrons is a linear function of the frequency of the radiation producing them. This leads to the existence of a universal constant with the dimensions of action, or of an elementary quantum of action ($h$); its value is found from measurements to be $6{\cdot}55 \times 10^{-27}$ erg. sec.

§ 91. The concept of individual light-quanta ($h\nu$) explains the fundamental relations to be observed in the reciprocal transformation between wave radiation and material radiation. The Compton effect shows that a momentum of magnitude $h\nu/c$ must be ascribed to light-quanta, and the displacement towards the red proves that they are subject to gravitational action. A complement to the corpuscular theory of light is formed by Louis de Broglie's theory of material waves, according to which every moving particle of matter is associated with a characteristic wave-length $h/mv$.

§ 92. Observations on the passage of $\alpha$-rays through matter show that an $\alpha$-particle can penetrate through several thousands of atoms without experiencing any appreciable change in direction, although sometimes a single atom produces a very large deflection of the $\alpha$-particle. The deflection must therefore be due to an atomic nucleus which occupies only a very small part of the total volume, and in which almost the whole mass of the atom is concentrated. According to Rutherford's fundamental conception, the nucleus must be regarded as positive and as surrounded by revolving electrons, the electrical force of attraction exerted on these being balanced by the centrifugal force.

§ 93. The method of electron impacts renders evident the existence of an excitation energy of atoms, and it allows of the experimental determination of characteristic potentials for many elements. In conjunction with the concept of light-quanta, the occurrence of energy-levels in atoms

affords an explanation for the appearance of characteristic atomic line spectra, for the coincidence of emission and absorption lines, and for the possibility of representing spectra by terms. The values experimentally found for the characteristic potentials agree with those calculated indirectly from spectroscopic data.

§ 94. As Moseley discovered, the elements can be arranged in a consecutively numbered series, in such a way that the square root of the frequency of a definite X-ray line increases linearly with the number (the atomic number) of the element. The "natural sequence" of the elements thus obtained, which begins with hydrogen and ends with uranium, comprises 92 places, only two of which are unoccupied. The atomic number proves to be identical with the nuclear charge number.

§ 95. The natural sequence of the elements is divided by the chemically inert rare gases into seven periods. The first period contains 2 elements, the second and third periods comprise 8 elements each, the fourth and fifth 18 each, the sixth contains 32 elements, and the seventh period breaks off at the sixth place with uranium. Homology among the elements allows of a distinction between eight groups, each with two sub-groups, in the periodic system.

## Chapter XII

### The Principles of Atomic Mechanics

§ 96. In its application to uniform motion in a closed orbit, the concept of material waves leads to the conclusion that the only possible orbits are those for which the length of path is an integral multiple of the wave-length, so that the angular momentum must be an integral multiple of $h/2\pi$. The velocity of material waves proves to be inversely proportional to the corresponding mechanical velocity, and it follows that the mechanical principle of Maupertuis is a necessary consequence of Fermat's wave principle. The mechanical velocity of a particle represents the group velocity of the material waves associated with it.

§ 97. Developing de Broglie's ideas, Schrödinger aimed at a generalization of mechanics which should be analogous to the transition from Fermat's to Huygens' principle, *i.e.*, from geometrical ray optics to physical wave optics. Its

application to a single particle led to the fundamental differential equation

$$\nabla^2 S + \frac{8\pi^2 m}{h^2}(E - V)S = 0,$$

where $S$ denotes the so-called field-scalar, $E$ the total, and $V$ the potential energy. By this relation the physical problem of quantizing energy is evidently reduced to the purely mathematical problem of determining the proper values of differential equations. The cases of the linear harmonic oscillator and of the rotator with a free axis are treated as examples. In contradistinction to the earlier theory, the energy of the oscillator is found to be equal to an odd multiple of half an energy element, and the hypothesis of the energy of the absolute zero becomes at once intelligible. The value obtained for the angular momentum of a rotator is $\frac{h}{2\pi}\sqrt{n(n+1)}$.

§ 98. Heisenberg replaced each periodically variable coordinate of an electron by a quadratic scheme of partial vibrations, the frequencies of which agree with those of the spectral lines, whilst the amplitudes determine the intensities of the lines. He defined multiplication for these schemes in such a way that the operation is not commutative. The difference resulting on an interchange of the factors in the product of a coordinate and the corresponding momentum is equated to $h/2\pi i$, by Heisenberg's commutation rule. Born and Jordan have facilitated an exact execution of Heisenberg's programme by the creation of matrix mechanics and, with the help of the latter, they were able to show that the Law of the Conservation of Energy and the Principle of Light-Quanta followed necessarily from Heisenberg's commutation rule. An application of the concept of the mathematical operator establishes the equivalence of Schrödinger's and Heisenberg's theories.

§ 99. Born interpreted Schrödinger's field-scalar as the probability of the coincidence of a material particle with a given point. It appears from quantum mechanics that only statistical probabilities such as this are definite; the elementary processes of physics are not so determined.

§ 100. Considerations of atomic mechanics lead to a distinction between the principle quantum number ($n$), which determines approximately the energy for an atom with one electron, and the subordinate quantum number ($l$), which defines the angular momentum of the periodic orbital

motion of the electron, and can assume the values 0, 1, 2 . . . $(n-1)$. Furthermore, a proper rotation, or so-called spin, must be ascribed to the electrons, with an associated quantum number $(\pm 1/2)$. According to a fundamental assumption of Russell and Saunders, the composition of the vectors determining the angular momentum and spin for the individual electrons in an atom must yield a resultant which is itself quantized.

§ 101. The resultant of the vectors of angular momentum and spin for the individual electrons defines the inner quantum number $(j)$ of the atom. The number of possible values for $j$ represents the multiplicity of the spectral terms: for $j = 0$ only singlet terms can exist. Within a multiplet, the frequency intervals obey simple numerical relations which are given by Landé's rule.

§ 102. In a magnetic field, a $(2j + 1)$-fold quantization of orientation exists for an atom, since the magnetic quantum number $(m)$ can assume all values from $(-j)$ to $(+j)$ differing by unity. In a weak field, the displacement of a term is equal to $g\, m\, \omega\, H$, where $H$ is the magnetic strength of field, $\omega$ a universal constant, and $g$ Landé's resolving factor. On the other hand, the displacement in a strong field is always an integral multiple of $\omega\, H$.

§ 103. Pauli's principle, according to which the values of all four quantum numbers can never be the same for any two electrons in an atom, gives, as the maximum number of electrons in a complete group, 2 for a one-quantum, 8 for a two-quantum, 18 for a three-quantum, and 32 for a four-quantum group.

## Chapter XIII

### Atomic Spectra

§ 104. As was substantially discovered by Balmer, the wave-numbers of the known hydrogen series can be represented as the differences between pairs of terms, which are equal to the quotient of the universal Rydberg's constant $(R)$ and the square of a whole number. An interpretation of this rule from the point of view of atomic theory yielded a value of $2\,\pi^2\, e^4\, m/h^3$ for Rydberg's constant. The fine structure of the hydrogen spectrum was explained first by Summerfeld, by introducing a relativistic correction, and subsequently on the basis of the electron spin.

§ 105. If in the formula for the hydrogen spectrum the

frequency be multiplied by four, the expression thus obtained must represent the spectrum of ionized helium. In point of fact these series, which were originally erroneously ascribed to hydrogen, have been shown by later experiments to be due to helium. The displacement towards the violet of the helium lines relatively to the hydrogen lines is explained, as Bohr recognized, by the simultaneous motion of the atomic nucleus. Measurements on this displacement yield an exact determination of the specific charge of an electron.

§ 106. The most important lines in the alkali spectra can be referred to four series of terms, which are called the *s*-, *p*-, *d*- and *f*-terms, and characterized by the subordinate quantum numbers 0, 1, 2, and 3. The principal series, the diffuse subordinate series, the sharp subordinate series, and the fundamental series correspond, in the case of emission, to the transitions $p \rightarrow s$, $d \rightarrow p$, $s \rightarrow p$, and $f \rightarrow d$. With the exception of the *s*-terms, which are always simple, the alkali terms are doublets, and they can be represented in the form $R/(n + a + a/n^2)^2$, where $a$ denotes the Rydberg correction (always negative) and $a$ the Ritz correction (always positive), both corrections depending only on the subordinate quantum number. The arc spectra of the alkali metals are analogous to the spark spectra of the alkaline earths, and to the higher spark spectra of atoms several times ionized, from which all the valency electrons except one have been stripped.

§ 107. Atoms with two valency electrons are found to possess a singlet and a triplet system of terms. Furthermore, displaced terms also occur; these are associated with a state in which two electrons are simultaneously excited. The arc spectrum of helium, consisting of two systems of terms between which combinations scarcely ever occur, may be regarded as a degenerate alkaline earth spectrum.

§ 108. The terms afforded by a system of equivalent electrons, *viz.*, by a system of electrons with the same principal and the same subordinate quantum numbers, may be calculated with the help of Pauli's principle. Addition of a further, non-equivalent, electron gives rise to new terms with a multiplicity larger or smaller by unity.

§ 109. The essential features of the spectra of elements in the short periods can be derived from results obtained with systems of equivalent electrons. Aluminium exhibits a simple system of doublets: in the sixth group triplet and quintet terms occur. The spark spectra of the alkali metals prove to be analogous to the arc spectra of the rare gases.

§ 110. A doubly ionized scandium atom contains a *d*-electron as a single three-quantum electron, and this fact explains the electronic grouping and fundamental terms characteristic of the long periods. The chemical ambiguity of copper has a spectroscopic parallel. General relations for the more complicated spectra find expression in Sommerfeld's alternation rule (governing the multiplicity), in Slater's fundamental-term rule, and in Laporte and Russell's combination rule.

§ 111. X-ray spectroscopy, as originated by Moseley, is based on Barkla's discovery of characteristic X-rays. The *K*-, *L*-, *M*-, and *N*-series arise from transitions to a one-, two-, three-, and four-quantum final state. Since the excitation potential necessary for the generation of a definite X-ray line is always found to be equal to the characteristic potential of the final level, an X-ray line can be produced obviously only when an electron in the emitting atom has been previously removed from its normal orbit to beyond the limits of the atom. This hypothesis is afforded a direct experimental confirmation by the absorption edges in the continuous X-ray spectrum. The scheme of X-ray terms appears closely related to the term system of the alkali metals, and, like the latter, is a doublet system in which *s*-, *p*-, *d*-, and *f*-terms occur. Terms differing only in the inner quantum number form regular doublets, the wave-length difference for which is almost independent of the atomic number of the element.

§ 112. A direct confirmation of hypotheses as to the completion of the inner groups of electrons is afforded by the course of the curves representing the relation between the square roots of the X-ray terms and the atomic numbers. The close similarity in chemical and physical behaviour between the rare earths is explained by the fact that the four-quantum group of electrons is not completed from eighteen to thirty-two electrons until in the sixth period. Bohr's theory of the periodic system indicated that the seventy-second element could not be a rare earth, and thus led to the discovery of hafnium. Elements containing incomplete sub-groups of electrons frequently exhibit the phenomena of paramagnetism, multiple valency, and the formation of coloured salts.

§ 113. The difference between the normal and the anomalous Zeeman effect is referred by the quantum theory to a distinction between strong and weak fields. The normal effect always yields an equidistant line triplet; the pheno-

mena of the anomalous effect can be correctly predicted on the basis of Landé's resolution formula. When a weak field is gradually strengthened, a transformation of the anomalous into the normal effect commences, giving rise to the Paschen-Back effect.

§ 114. The resolution of spectral lines in an electric field, or the so-called Stark effect, may be regarded as an experimental confirmation of the quantum theory.

§ 115. Simple, rational, numerical ratios connect the intensities of the components of a line multiplet. It is found that they obey accurately a relation which states that the sum of the intensities of the lines, having a common term as either initial or final term, is proportional to the weight of this term, as defined by the number $(2j + 1)$. For a triplet in the diffuse subordinate series the intensity ratio is 9 : 5 : 1, the central line being the strongest, and the line of largest wave-length the weakest. With the doublets in the principal series, the component of smaller wave-length is twice as strong as the other; in the sharp subordinate series the converse holds good.

§ 116. Researches with atomic rays carried out by Gerlach and Stern afforded a direct experimental proof for the quantization of orientation of atoms in a magnetic field demanded by theory. It was established that a silver atom takes up such a position that the component of its magnetic moment amounts exactly to a magneton, which is defined as the moment $(e\,h/4\,\pi\,m\,c)$ corresponding to a mechanical angular momentum $h/2\,\pi$. In accord with theoretical expectations, the atomic rays of those elements, for which the fundamental term is one with $j = 0$, are not resolved. Langevin's theory of the temperature variation of paramagnetism has been combined with the new concept of quantization of direction to produce a formula which, satisfactorily confirmed by experiment, connects the magnetic, so-called Curie, constant with the inner quantum number and Landé's resolving factor.

§ 117. The excitation of atoms by an access of translational energy manifests itself as thermal luminescence. Ionization to the same degree demands higher temperatures than excitation, and in this lies the explanation of the earlier paradoxical results obtained in solar spectroscopy. In the interior of stars, atoms are largely deprived of their planetary electrons, so that the enormous density of many stars becomes intelligible. Fluorescence in gases can be extinguished by increasing the pressure; observations made in

this connection have shown that the value of the mean duration of an excited atomic state is about $10^{-8}$ sec. An excited atomic state, from which a spontaneous return to the normal state appears impossible, in view of the spectroscopic selection principles, is termed metastable. Metastability has been observed principally with helium, but it has also been detected with neon, argon, and other elements. The so-called nebulium, lines of which occur in the spectra of cosmic nebulæ, proves to be metastable oxygen and nitrogen.

## Chapter XIV

### X-Rays

§ 118. A space-lattice separates by diffraction rays of quite definite wave-lengths from out of an incident wave radiation. These wave-lengths depend, for given ordinal numbers, on the lattice-constant and the direction of incidence.

§ 119. The diffraction of rays by a space-lattice may also be regarded as a reflection at the lattice-planes. A set of lattice-planes singles out from an incident radiation those rays for which an integral multiple of the wave-length is equal to the difference in path between two rays that are reflected from two successive lattice-planes. On rotation of a space-lattice, all rays of the same wave-lengths are focussed at definite points lying on the circumference of a circle.

§ 120. The diffraction of X-rays by crystals was demonstrated by Laue. The investigation of the structure of a crystal on the basis of the Laue photographs rendered possible an exact measurement of the wave-lengths of X-rays. By means of crystals, wave-lengths of from about 50 to 50,000 X can be measured.

§ 121. Electron rays exhibit properties similar to those of X-rays. They are selectively reflected by crystals, as was discovered by Davisson and Germer in an application of de Broglie's theory, and, as G. P. Thomson and Rupp have shown, they are diffracted on a passage through foils.

## Chapter XV

### The Atomic Nucleus

§ 122. The phenomenon of isotopy, *viz.*, the apparent chemical identity of substances with different radioactive properties or atomic weights, is explained by the fact that

two atoms with the same nuclear charge may yet possess a different nuclear structure, and thus a different number of protons. By means of Aston's mass-spectroscopy, the majority of elements have already been investigated for isotopy, and a separation of isotopes has also been experimentally effected, although only to a very small degree.

§ 123. In accordance with the Law of Mass of Energy, the small deviations from integral values exhibited by the mass numbers of the individual atomic species are explained by the so-called packing effect. The deviation (referred to oxygen) divided by the number of protons gives the packing ratio ; this is largest for hydrogen, positive up to neon, and then negative on to mercury, when it becomes again positive.

§ 124. The phenomena of radioactivity are due to a nuclear disintegration which, owing to the alteration in the nuclear charge number, at the same time effects a chemical transformation. An $\alpha$-transformation reduces the atomic number by two, and the atomic weight by four, whereas a $\beta$-transformation raises the atomic number by one, without noticeably altering the atomic weight. The observation of the rate of formation of helium from radium, combined with the scintillation method, permits of a direct determination of the Loschmidt number.

§ 125. In radioactive equilibrium the numbers of atoms present are in inverse ratio to the disintegration constants. Equal numbers of atoms of all substances which are in equilibrium with each other disintegrate in the same time.

§ 126. At the present time three transformation series are known, commencing with uranium, actinium, and thorium respectively, and all ending with stable species of lead. Besides these series, potassium and rubidium are also known to be radioelements. The age of uranium minerals can be calculated from their content of uranium lead.

§ 127. The range of $\alpha$-rays is proportional to the cube of their initial velocity. The logarithms of the disintegration constants are linear functions of the logarithms of the ranges. $\beta$-ray spectra are produced by the deflection of $\beta$-rays in magnetic fields.

§ 128. The wave-lengths of $\gamma$-rays can be determined in part directly, by means of interference phenomena, and in part indirectly from magnetic $\beta$-ray spectra. Similar combinational relations are found to hold for $\gamma$-lines as for optical or X-ray lines. The number of $\gamma$-light-quanta emitted proves to be equal to the number of the disintegrating atoms.

§ 129. Whereas $\gamma$-rays are equivalent to potentials up to

2,000,000 volts, the cosmic rays reaching the earth from space correspond to potentials extending to over 200,000,000 volts. Millikan has found that cosmic radiation is composed of groups of different wave-lengths, which may possibly be ascribed to the formation of helium, oxygen, and silicon nuclei from protons and electrons.

§ 130. By means of α-rays, the nuclei of numerous elements have been disrupted, hydrogen nuclei being expelled in the process.

§ 131. Elements of even atomic number are of more frequent occurrence than those of odd number. For the large majority of atomic species the number of nuclear electrons is even. Atomic species with an atomic weight divisible by four are particularly common.

## Chapter XVI

### The Molecule

§ 132. The principal relations of molecular spectra can be explained on the hypothesis that the molecular energy consists of three components, each of which can assume only discrete, stepped, values. These components are the rotational energy, the energy of the nuclear vibrations, and the energy of the electrons. When the electronic energy remains constant on a quantum transition, a rotation-vibration spectrum is produced ; this is composed of bands of equidistant lines, with the zero-lines missing.

§ 133. Within a band spectrum, every electron transition is associated with a band system consisting of separate bands, the zero positions in which are connected by the band-head law. In general each band is made up of three branches, corresponding to the three possible changes in the rotational quantum number. As regards the electron terms, a far-reaching analogy can be established between molecular and atomic spectra, on the assumption that the atoms forming a molecule possess common shells of electrons. For example, the spectra of molecules with thirteen or twenty-one electrons resemble the atomic spectra of the alkali metals. Diatomic elementary molecules exhibit an alternation of intensity in their bands—an effect recognized by Heisenberg as due to a quantum-mechanical resonance.

§ 134. From the band spectra of diatomic molecules the molecular moment of inertia, and hence the nuclear interval, can be calculated. The nuclear intervals and nuclear

frequencies of molecules are evidently related to the chemical periodicity. The convergence points of band spectra offer the possibility of an optical determination of the heat of dissociation. Dissociation may be effected by the absorption of light-quanta, sometimes with the help of sensitization.

§ 135. The potential relations in molecular compounds show that the assumption of electrostatic forces is not sufficient. On this basis of wave mechanics, Heitler and London obtained two different solutions for the energy of interaction between two neutral hydrogen atoms in the normal state ; one of these solutions indicates a homopolar attraction and the other a continuous repulsion. The former solution yields values, which agree with experiment, for the nuclear interval and the dissociation energy of the hydrogen molecule. For two neutral helium atoms in the normal state, only the solution indicative of repulsion is apparently possible. Quite generally, according to London, homopolar compounds appear, on the basis of Pauli's principle, as a pairing of two electrons which, with the same orbital motions, belong as unpaired electrons to different atoms. The number of unpaired electrons in an atom determine its chemical valency. In this way the peculiar positions of oxygen and of fluorine can be explained.

§ 136. The lattice theory of crystals leads to relations between elastic and optical constants of crystals ; between their optical constants and the melting-point temperature ; between cohesion and chemical forces. It affords also a basis for a theory of thermal conductivity, of optical double refraction, of lattice types, and of crystal growth.

§ 137. Debye's theory of electrolytes rests on the assumption of ionic clouds which surround the ions in very dilute solutions, and require a certain " time of relaxation " for their development. In this way Debye's theory explains the so-called Wien effect, which consists in the increase in conductivity with the potential in strong fields.

§ 138. The electric moment produced in a dielectric is made up of a deformation component independent of the temperature, and of an orientation component, which varies with the temperature, but exists only with polar molecules. Hence the dielectric constant is independent of the temperature only in symmetrical molecules, whereas the temperature effect otherwise observed allows of a calculation of the polar moment.

§ 139. The characteristics of metals can be explained by the assumption of free electrons. Sommerfeld applied to

the latter Fermi's statistics, which are based on Pauli's principle. Because of the insignificance of their mass, free electrons represent a completely degenerate gas, even at room temperatures. This accounts for the fact, inexplicable by the older theory, that free electrons do not contribute measurably to the specific heat of metals.

## Chapter XVII

### The Interaction between Light and Matter

§ 140. According to the new quantum theory, the phenomena of dispersion depend, not on the actual oscillational or orbital frequencies of the electrons in the atoms, but on the whole of the possible spectral frequencies of the atom struck by the light-wave, although these frequencies need not actually occur. The conclusion of the occurrence of combinational frequencies is apparently confirmed by the Raman effect.

§ 141. The dispersion theory of wave mechanics also affords an explanation for the photo-effect, which is apparently determined by a material wave emitted by the atom concerned. The Auger effect is to be regarded as a special form of the photo-effect.

§ 142. On a transformation of cathode rays into X-rays, it is found that the light-quantum relation is exactly obeyed.

§ 143. It follows from the Laws of the Conservation of Momentum and of Energy that, on a collision between a light-quantum and a free electron, the wave-length changes by an amount which depends only on the angle of scatter, but not on the primary wave-length. This phenomenon is actually observed in the Compton effect. Experiment shows that the two laws of conservation hold, not only statistically, but also individually. A wave-mechanical theory for the Compton effect has been developed by Wentzel, taking into account the association of the scattering electron with the atom. In this way the intensity ratio between displaced and undisplaced lines has been derived.

§ 144. Dirac has succeeded in establishing equations which bear the same relation to the wave equation of matter as do Maxwell's equations of the electromagnetic field to the wave equation of light. On a relativistic basis, Dirac's equations of themselves account for the phenomenon of electronic spin.

§ 145. Several astronomical methods agree in assigning

an age of 5 to 10 $\times$ $10^{12}$ years to the stars and the sun. The helium synthesis, however, as the strongest source of atomic energy, could only provide about two thousandths of the energy requirements of the sun over this long period. On the other hand, a dissolution of solar matter (in the sense of the theorem of proper energy) would suffice for the present rate of emission of energy by the sun, already diminished to two ergs/gram sec., for 15 $\times$ $10^{12}$ years.

## Part IV

# THEORY OF HEAT

## Chapter XVIII

### Statistics

§ 146. The consideration of individual distributions of state leads to the concept of statistical probability; by this is understood the number of complexions through which a definite distribution can be realized. If $w_i$ be the fraction of individuals in the state characterized by the index $i$, and $N$ be their total number, the probability logarithm is given by the relation

$$\log_e W = -N \sum w_i \log_e w_i .$$

It follows directly from the concept of statistical probability that the transition between two distributions in the direction of increasing statistical probability is more probable than that in the reverse direction. On the other hand, since there must be an upper limit for the statistical probability, every system in which the distribution of state is constantly changing must exhibit an apparent tendency to assume the most probable distribution and, when this has been attained, to persist in it.

§ 147. Let a system possess a definite amount of a certain quantity, this amount being equal to the sum of the component amounts associated with the separate individuals. Moreover, let the component amount associated with an individual in a definite individual state be completely determined by quantities which, at any moment, have the same value for all individuals in the system and are therefore called parameters of the system. A distribution of which the statistical probability represents an extreme value for those variations of the distribution wherein the

total amount, the number of individuals, and the parameters remain unaltered, is termed canonical. It may be shown that a canonical distribution is the most probable one, every system tending to assume it. Infinitesimal canonical variations occur without any change in the statistical probability, and thus appear to be equally probable in either direction; they are apparently reversible.

§ 148. If the quantity previously considered be identified with an amount of energy ($E$), two important formulæ result. If $dQ$ be an infinitesimal amount of energy introduced into a system with the energy $E$, whilst the work $dA$ is simultaneously performed by the variation of its parameters, then

$$dQ = dE + dA .$$

Furthermore,

$$\frac{dQ}{\Theta} = d\,(\log_e W),$$

where $\Theta$ is the modulus of distribution of the energy. Thus, on the one hand the difference between the energy introduced and the work performed, and on the other hand the quotient of the energy introduced and the modulus, represent the complete differentials of two functions which depend only on the instantaneous state of the system, being completely determined by the modulus and the parameters. In accordance with a previous statement, the second function has the property of tending towards a maximum, and of remaining constant only in reversible processes.

§ 149. The probability logarithm of a composite system is equal to the sum of the probability logarithms of the component systems. It represents a maximum when all the component systems are canonically distributed and have the same modulus, and, in that case, statistical equilibrium is said to exist. When the values of the modulus are different there is a tendency towards equalization, in such a way that energy is not spontaneously transferred from a system with a smaller to one with a larger value of the modulus.

§ 150. The negative partial derivatives of the free energy with respect to the parameters may be defined as statistical forces, and, together with the parameters and the modulus of distribution, they may be regarded as quantities of state of the system. It may be shown that $r$ equations of state must exist between $(2r + 1)$ quantities of state, so that, of the latter, any arbitrary $(r + 1)$ can be selected as independent variables of state.

§ 151. In a consideration of cyclic processes consisting of

two isomodular and two adiabatic component processes, the ratio between the work performed and the energy introduced during the first isomodular component process may be regarded as the efficiency of the cyclic process. Measurements of the efficiency enable an observer to prepare an exact scale of modulus values, even when the interior of the system is completely hidden from him and he can only follow the behaviour of the system from outside.

§ 152. The mean value of the individual energy of a system of particles, the energy of which is purely kinetic, proves to be three times as large as half the modulus of distribution: when there are $s$ degrees of freedom, it is $s$ times as large. This so-called theorem of equipartition allows of the derivation of Maxwell's law of the distribution of velocities, with the help of the principle of the canonical distribution of energy.

§ 153. It can be shown, with the aid of the theorem of the virial, that, in the case of a system of particles with purely kinetic energy, the product of the pressure and the volume is equal to the product of the number of particles and the modulus of distribution.

§ 154. According to Rayleigh and Jeans's Principle, in a system of stationary vibrations the density of the characteristic vibrations with frequencies lying in the interval between $\nu$ and $(\nu + d\nu)$ is equal to $4\pi\nu^2/v^3$ ($v$ being the velocity of propagation of the vibrations).

§ 155. A system of stationary vibrations is in statistical equilibrium when all the component systems of different frequencies have the same value of the modulus of distribution. In a quantized vibrating system, the energy density is proportional to the fourth power of the modulus of distribution.

## Chapter XIX

### Thermodynamics

§ 156. The energy distributed amongst the molecules of a body is interpreted in physics as the heat content of the body. In those cases in which energy apparently disappears whilst heat is evolved, this heat can be directly measured in energy units. Moreover, by carrying out cyclic processes and measuring their efficiencies, an exact scale of values of heat moduli can be established. The usual temperature scale

represents a conventional and arbitrary scale of such values.

§ 157. The establishment of a temperature scale further permits of the creation of an arbitrary unit of quantity of heat. The energy, measured in energy units, which is equivalent to this arbitrary unit of heat, represents the mechanical equivalent of heat.

§ 158. The theorems derived in general statistics must still hold good if, in them, modulus of distribution be made proportional to the temperature, and heat be substituted for energy. In this way the two Principal Laws of Thermodynamics are obtained: I. If an infinitesimal quantity of heat be introduced into a system, while work is simultaneously performed, the difference between heat introduced and work performed represents the complete differential of a function, the internal energy, which depends only on the instantaneous state of the system and, when $r$ parameters exist, is therefore determined by $(r + 1)$ independent variables of state. II. The quotient of the infinitesimal quantity of heat introduced and the temperature,

$$\frac{dQ}{T} = dS\,,$$

represents the complete differential of a second function, known as entropy, which likewise depends only on the instantaneous state of the system, and is therefore also completely determined by $(r + 1)$ independent variables of state. If the ratio between modulus of heat and temperature be denoted by the universal constant $k$, *i.e.*, if

$$\Theta = k\,T\,,$$

then the entropy

$$S = k \log_e W$$

(Boltzmann's entropy law). Entropy must possess the same properties as the probability logarithm. Entropy appears to tend to a maximum, and to remain constant only in reversible processes. Different bodies can only be in thermodynamic equilibrium when their temperatures are the same. When their temperatures differ, there is a tendency to equalization, in such a way that heat is not of itself transferred from a colder to a warmer body.

§ 159. The law of the increase of entropy is no law in the strict sense of the word, but only a rule which is far more likely to be obeyed than disobeyed. On account of the smallness of the molecules, noticeable exceptions are rare

beyond any human power of imagination, and may hence be regarded as practically excluded.

§ 160. In isothermal processes, the work performed is equal to the decrease in free energy. Again, to processes in which no external work is performed Helmholtz's equation,

$$E - F = -T\frac{dF}{dT},$$

applies. The concept of free energy ($F$) is closely connected with that of thermodynamic potential which, for a homogeneous system, assumes the form

$$\Phi = E - TS + pV.$$

§ 161. The coefficients of expansion, pressure, and compressibility, which characterize the changes of state of a system, are interconnected by an equation from which, if two of these quantities be known, the third can be calculated. Since both internal energy and entropy can be expressed either as functions of the temperature and the volume, or as functions of the temperature and the pressure, a series of theorems on the changes of state of a homogeneous system can be derived. For example, for an isothermal increase in volume, a body absorbs or emits heat according as its pressure coefficient at the temperature in question is positive or negative. If the expansion coefficient be positive, an adiabatic increase in pressure produces a heating effect, and an adiabatic increase in tension results in a cooling effect. If the expansion coefficient be negative, these effects are reversed.

§ 162. In Thomson and Joule's experiment, gas is allowed to flow adiabatically, under a pressure head, from one vessel into another. The experimental results permit of the reduction of the scale of a thermometer to the thermodynamic temperature scale.

§ 163. A system consisting of different phases of a substance in thermodynamic equilibrium represents a saturated complex of coexisting phases. Equilibrium is only possible when all the coexisting phases are at the same pressure and the same temperature, and when unit-masses of the phases have the same thermodynamic potential. Hence, when two phases coexist, there must be a functional relation between saturation pressure and saturation temperature. In a graphical representation in a $p$-$T$-plane, a saturation curve corresponds to this relation, and in the case of three coexisting phases, a so-called triple-point.

§ 164. For the thermodynamical equilibrium of a heterogeneous system, it is found as a necessary condition that, not only the temperature and the pressure, but also the Gibbs potentials of the several components must have the same values throughout the individual phases. The Gibbs potentials are the partial derivatives of the energy with respect to the individual masses at constant entropy and constant volume; they represent generalizations of the thermodynamic potentials referred to unit-mass. The number ($r$) of coexisting phases can never exceed the number of components ($n$) by more than two. The number of degrees of freedom of the system is then equal to $(n + 2 - r)$. These two theorems constitute the phase rule of Gibbs.

§ 165. In a system of molecules, the energy of which is purely kinetic, the equation of state assumes the simple form

$$p\,v = R\,T\,,$$

where $v$ is the specific volume referred to a gram-molecule, and $R$ is the product of the constant $k$ and the Loschmidt number. This equation represents the ideal gas law. The hypotheses leading to this equation are satisfied when the gas molecules are imagined as executing rapid translatory motions, the distances between them, however, being so large that any eventual forces acting between the molecules can be neglected. It follows from the ideal gas law that, at constant temperature, pressure and volume are inversely proportional to each other (Boyle's Law); furthermore, that at the same pressure and temperature, equal volumes of different gases must contain the same number of molecules (Avogadro's Law). The ratio of the two specific heats is equal to 5/3 for monatomic gases, whilst for polyatomic gases it must lie between 1 and 5/3.

§ 166. The molecules of a substance dissolved in a liquid exercise a so-called osmotic pressure which obeys the gas laws, and, with the help of semipermeable partitions, may be experimentally detected and measured. It may be indirectly derived from observations on the lowering of the vapour pressure, or from observations on the raising of the boiling-point or on the lowering of the freezing-point of solutions. The lowering of the vapour pressure stands in the same ratio to the vapour pressure itself as does the number of gram-molecules (or molecules) of the solute to those of the solvent. The changes in the boiling-point and the freezing-point are also proportional to the concentration.

§ 167. By taking into account the forces of attraction

between the molecules, and also the lower limit to the specific volume of a gas consequent on its molecular structure, van der Waals obtained the equation of state named after him :

$$\left(p + \frac{a}{v^2}\right)(v - b) = R\,T\,,$$

where $a$ and $b$ are constants. Provided that $T$ lies below a definite critical value, the isotherms represented by this equation have three real roots for $v$ ($v_1 < v_2 < v_3$), for a given value of $p$. According to van der Waals's theory, for any specific volume smaller than $v_1$ the substance must be liquid, whilst for any specific volume larger than $v_3$ it must be gaseous ; for values between $v_1$ and $v_3$ in general, however, the isotherms are to be replaced by straight lines parallel to the $v$-axis, corresponding to the coexistence of two states of aggregation. Hence, above the critical temperature, liquefaction of the gas is evidently impossible.

§ 168. The velocity of reaction in a chemical synthesis is proportional to the number of collisions between the molecules of the reacting species, and therefore to the product of the concentrations of the latter. Similarly, the velocity of decomposition of a substance is proportional to the square of its concentration. A generalization of these two theorems yields the chemical Law of Mass Action.

§ 169. As Einstein and Smoluchowski have shown, the Brownian movement is to be referred to the resultant displacements, suffered by particles suspended in a liquid, within such intervals of time as can be distinguished by the observer. The mean square of the displacement is proportional to that interval ; its experimental determination allows of an exact calculation of the Loschmidt number. A simple relation, derived by Einstein, connects the mean square of the displacement with the coefficient of diffusion.

§ 170. In suspensions, emulsions, and colloidal solutions, a so-called sedimentation equilibrium is formed under the dual influence of the osmotic pressure and the weight of the suspended particles ; investigation of this equilibrium leads to a direct, very accurate, determination of the Loschmidt number. The relative spontaneous fluctuations in the concentration are inversely proportional to the square root of the concentration. The blue colour of the sky is to be referred to molecular fluctuations in the density of the atmosphere and the consequent optical opalescence. Measurements of the intensity ratio between the blue of the sky and

sunlight also allow of a determination of the Loschmidt number.

## Chapter XX

### The Quantum Theory of Heat

§ 171. Modern theory refers the heat content of solid bodies to molecular oscillations, the energy of which, in the sense of the quantum hypothesis, is composed of energy-elements of magnitude $h\nu$. An approximate derivation leads to Dulong and Petit's Law, which has long been known, and according to which the specific heat referred to a gram-atom, or the so-called atomic heat, is equal to three times the gas-constant $R$, or to about six calories. The more modern theory, however, also affords an explanation for the deviations from Dulong and Petit's Law, which were discovered at normal temperatures in the case of diamond and are observed at low temperatures for all bodies.

§ 172. If, as Debye has suggested, solid bodies are to be regarded approximately as continua filled with stationary elastic waves, statistical theory leads to the conclusion that, at very low temperatures, the specific heat is proportional to the cube of the temperature (Debye's Law), whilst, at the absolute zero, both the temperature coefficient of the free energy and the entropy vanish (Nernst's heat theorem). Hence, at very low temperatures the expansion coefficient must also in general vanish.

§ 173. The entropy constant and the closely related chemical constant can be experimentally derived, on the basis of Nernst's heat theorem, from observations on the temperature variation of the heat of evaporation at very low temperatures. On the other hand, the concept of material waves, in conjunction with Rayleigh and Jeans's principle, leads to the conclusion that we must suppose the momentum space of a monatomic gas to be divided into cells which are equal in volume to $h^3$ divided by the volume of the gas. In this way, with the help of considerations of general statistics, we obtain a definite expression for the chemical constant. A comparison of the theoretical results thus derived with the experimental values yields a satisfactory agreement.

§ 174. A system of stationary electromagnetic vibrations in thermodynamic equilibrium with any bodies is known as cavity radiation. The energy density of cavity radiation is proportional to the fourth power of the temperature, as is also the emissive power of a black body (Stefan's Law),

because, as a geometrical consideration shows, the emissive power is equal to one-quarter of the energy density of the radiation, multiplied by the velocity of light.

§ 175. A spectral resolution of the emissive power leads to Planck's Law of Radiation, which gives the relation between the specific emissive power and the colour and temperature. The specific emissive power possesses a maximum for a certain wave-length, the displacement of which with the temperature is determined by the relation

$$\lambda_{\max} T = b$$

(Wien's Displacement Law). It appears from the quantum theory that the universal constant $b$ is equal to $c\,h/\beta\,k$, where $\beta$ (the root of a transcendental equation) represents a pure number nearly equal to 5. From the constants of Stefan's and Wien's Laws the values of $h$ and $k$ can be calculated.

# Part V

# THEORY OF RELATIVITY

## Chapter XXI

### Special Theory of Relativity

§ 176. Classical physics appeared to demand the existence of only a single reference system in which light is propagated in all directions with the same velocity. The possibility was thus presented of detecting an absolute motion of the earth by a suitable optical experiment, such as was designed by Michelson. The actual execution of this experiment, however, showed that the expected effect definitely did not exist.

§ 177. The apparent contradiction between theoretical physics and Michelson's experiment was solved by Einstein, when he showed that the prejudice of an absolute time was untenable. Einstein's Principle of Relativity maintains that the statements of time, in terms of which a physical process is described, are relative and depend on the point of view of the observer using them, whilst they are governed by the postulate that, when measured in terms of them, light shall appear to the observer to travel with the same velocity in every direction.

§ 178. According to Minkowski, Einstein's Principle of

Relativity may be geometrically interpreted by stating that the square of the product of the time and the velocity of light, taken with a negative sign, and the squares of the three spatial coordinates stand in the same relation to each other as do the squares of the four coordinates in a four-dimensional geometry. The four-dimensional manifold, which appears as a linking together of space and time by the coordinates $x$, $y$, $z$, and $i\,c\,t$, is known as the Minkowski-world. The motion of a particle is represented by a world-line, the directional cosines of which with respect to a definite four-dimensional system of coordinates determine the velocity of the particle relatively to the reference system used.

§ 179. The transition from one spatial coordinate system to another which is moving uniformly in the common $x$-direction with a velocity $v$ relatively to the first system appears, in a four-dimensional representation, as a rotation of the coordinate axes in the $x$-$t$ plane ; the tangent of the angle of rotation is equal to the imaginary quotient of the velocity $v$ and the velocity of light. From this and from the general formulæ for the transformation of vector components follow the equations expressing a Lorentz-transformation :

$$x' = \frac{x - v\,t}{\sqrt{1 - \frac{v^2}{c^2}}}, \quad y' = y, \quad z' = z, \quad t' = \frac{-\frac{v}{c^2}\,x + t}{\sqrt{1 - \frac{v^2}{c^2}}}.$$

§ 180. In consequence of the Principle of Relativity, the idea of simultaneity loses its absolute meaning. Two events which are apparently simultaneous to one observer do not appear simultaneous to another observer moving relatively to the first one. A moving body experiences a contraction in the direction of its motion, in the ratio $\sqrt{1 - v^2/c^2} : 1$. A " moving clock " goes more slowly than one at rest.

§ 181. According to the theory of relativity, the resultant velocity obtained by compounding two similarly directed velocities is given by

$$u = \frac{v + v'}{1 + \frac{v\,v'}{c^2}}.$$

It appears from this that velocities larger than that of light are impossible.

§ 182. It follows from the relativistic formula for the composition of velocities that light-waves partake in the

velocity of motion ($v$) of a flowing liquid to an extent given by the fraction $\alpha v$, where $\alpha$ is equal to $(1 - 1/n^2)$, $n$ being the refractive index of the liquid (Fizeau's experiment).

§ 183. Relativistic dynamics shows that the momental mass ($m$) and the stationary mass ($\mu$) are connected by the relation $m = \mu/\sqrt{1 - v^2/c^2}$. Force and acceleration are, in general, not similarly directed, but the direction of the acceleration includes a larger angle with the direction of motion than does the direction of the force.

§ 184. It follows from the theory of relativity that all energy possesses an inertial mass, which is equal to the quotient of the quantity of energy and the square of the velocity of light.

§ 185. In a four-dimensional manifold, skew-symmetrical tensors correspond to the axial vectors of three-dimensional geometry. The vector divergence of a skew-symmetrical tensor assumes a simple form, and the components of a skew-symmetrical tensor which results from the rotation of a world-vector are also interconnected by simple relations.

§ 186. In a four-dimensional representation, Maxwell's equations may be summarized in the statement that the four-current represents the vector divergence of the electromagnetic field tensor, which in its turn represents the rotation of a world-vector. The four-current can be resolved into convection current and charge, and the field tensor into electric and magnetic field-strengths.

§ 187. The product of mass density and the expression $(1 - v^2/c^2)$ represents a scalar. On multiplying the products (taken in pairs) of the components of the world-line-element representing the motion by this scalar, the components of the matter tensor are obtained. The space-time-components of this tensor represent the density of momentum, and its purely temporal component the density of matter.

## Chapter XXII

### Theory of Gravitation

§ 188. The general theory of relativity bases the description of the Minkowski-world, not on Euclidean, but on the more general Riemannian geometry, for it assumes that the Minkowski-world is "curved" in consequence of the space-time distribution of matter. The assertion that, in this

curved world, the motion of a particle upon which no electromagnetic forces are acting is to be described by a geodesic world-line represents both the generalization of the Law of Inertia, and also the new Law of Gravitation. From this concept follows at once the theorem of the identity of gravitational and inertial mass. In the general theory of relativity, Cartesian are replaced by the more general Gaussian coordinate systems, and thus any transition between two reference systems moving quite arbitrarily with respect to each other may be represented by a four-dimensional coordinate transformation.

§ 189. In Riemannian geometry, the Theorem of Pythagoras is replaced by the generalized formula

$$ds^2 = \sum_h \sum_k g_{hk}\, dx_h\, dx_k\,.$$

For a rectangular, Cartesian, coordinate system, the $g_{hk}$ are equal to unity or zero, according as $h$ is or is not equal to $k$.

§ 190. In utilizing Riemannian geometry, a distinction must be drawn between covariant and contravariant vectors, which are transformed differently on a change of coordinate system. Quantities, the components of which are transformed in the same way as the products formed by multiplying in pairs (or triplets, *etc.*) the components of vectors, are known as covariant, contravariant, or mixed tensors. The operation of contraction permits of the derivation from any mixed tensor of another tensor of a rank lower by two.

§ 191. The quantities $g_{hk}$ considered in § 189 are the components of a covariant tensor of the second rank, which is termed the metric fundamental tensor. If the determinant of the covariant quantities $g_{hk}$ be formed, the cofactors divided by the value of the determinant represent the components ($g^{hk}$) of the contravariant fundamental tensor.

§ 192. The equation of a geodesic in Riemannian geometry follows from its definition as the shortest line connecting two points. This equation contains Christoffel's three-indices symbols, which are formed of the components of the metric fundamental tensor and its partial derivatives with respect to the coordinates.

§ 193. Mathematical considerations based on the equation of a geodesic lead to a mixed tensor of the fourth rank, known as the Riemann tensor, which depends only on the three-indices symbols and their derivatives with respect to the coordinates. Contraction of this tensor yields a tensor

of the second rank (the contracted Riemann tensor $R_{hk}$), and from this can be obtained a scalar

$$R = \sum_h \sum_k g^{hk} R_{hk},$$

which is called the Riemann invariant.

§ 194. If the Riemann invariant be integrated over an arbitrarily delimited region of a geometrical manifold, and the variation of this integral be then formed in such a way that the fundamental tensor is varied only within the boundaries of this region, this process leads to a tensor

$$S_{hk} = R_{hk} - \tfrac{1}{2} g_{hk} R.$$

In Einstein's field equations, this tensor is equated to the matter tensor, apart from a universal proportionality factor ($f$), so that

$$R_{hk} - \tfrac{1}{2} g_{hk} R = f T_{hk}.$$

In the Minkowski-world, the curvature of which is described by these field equations, it is postulated that the motion of a particle shall be represented by a geodesic world-line. On the special assumptions that, first, the deviations from Euclidean geometry are small ; secondly, that the mechanical velocities involved are small as compared with the velocity of light ; and thirdly, that the gravitational field is static, Newton's Law of Gravitation may be derived as a first approximation. The gravitational constant ($\kappa$) is found to be equal to $f c^2/8\pi$.

§ 195. In a gravitational field produced by a stationary mass, all processes are apparently retarded in the ratio $1 : (1 + \alpha/r)$, where $\alpha$ denotes the gravitational radius of the mass responsible for the field, this radius being obtained on multiplying the mass by $\kappa/c^2$. Rays of light must be curved in a gravitational field, and also Kepler's Laws must assume a more complicated form than in the classical theory, owing to the necessary slow advance of the perihelion of a planetary orbit.

§ 196. Einstein's gravitational theory received apparent confirmation from three experimental facts : *viz.*, the perihelion motion of the planet Mercury ; the deflection of rays of light coming from stars on passing close to the sun, which has been detected and measured in total solar eclipses ; and, finally, the displacement towards the red of the spectral lines of stars.

§ 197. An extension of his Law of Gravitation has led Einstein to the assumption that space is spherically curved,

and thus, despite its boundlessness, that it possesses a finite volume. According to this view, the radius of curvature of the universe is equal to its gravitational radius divided by the numerical factor $\pi/2$. Astronomical estimates give a figure of about 100 to 100,000 million light-years for the extent of the universe, and a value of between $10^{54}$ and $10^{57}$ grams for its total mass.

# SYNOPSIS OF THE SYMBOLS MOST FREQUENTLY USED

## Part III

## ATOMIC THEORY

$E$ Energy, or electric field-strength.
$G$ Momentum.
$H$ Magnetic field-strength.
$I$ Molecular moment of inertia.
$L$ Loschmidt number, or kinetic energy.
$R$ Rydberg term.
$S$ Schrödinger's field-scalar.
$T$ Spectral term, or absolute temperature.
$U$ Angular momentum.
$V$ Potential energy or potential.

$a$ Radius, or lattice-constant.
$c$ Velocity of light.
$d$ Lattice-plane distance.
$e$ Elementary quantum of electricity.
$g$ Resolving factor, or group velocity.
$h$ Elementary quantum of action.
$j$ Inner quantum number.
$k$ Boltzmann's constant.
$l$ Subordinate quantum number.
$m$ Mass, or magnetic, or rotational quantum number.
$n$ Principal quantum number.
$p$ Momentum component.
$q$ Positional coordinate.
$r$ Distance, or quantum number of electron spin.
$t$ Time.
$v$ Velocity.

$\alpha, \beta, \gamma$ Directional cosines.
$\gamma$ Specific charge.
$\lambda$ Wave-length, or disintegration constant.
$\nu$ Frequency.

## Part IV

## THEORY OF HEAT

$A$ Work.
$C$ Molecular or atomic heat.
$D$ Coefficient of diffusion.
$E$ Energy.
$F$ Free energy.
$J$ Mechanical equivalent of heat.
$K$ Emissive power.
$L$ Loschmidt number.
$N$ Number of particles.
$P$ Osmotic pressure.
$Q$ Quantity of energy or heat.
$R$ Gas constant.
$S$ Entropy.
$T$ Absolute temperature.
$V$ Volume.
$W$ Statistical probability.

$a$ Stefan's constant, or the first of van der Waals's constants.
$b$ Wien's constant, or the second of van der Waals's constants.
$c$ Specific heat, or velocity of light.
$h$ Elementary quantum of action.
$i$ Chemical constant.
$k$ Boltzmann's constant.
$m$ Mass.
$n$ Number of particles in unit-volume, or of gram-molecules.
$p$ Pressure.
$q$ Specific quantity of heat.
$s$ Specific entropy.
$v$ Specific volume.

$\Theta$ Modulus of distribution.
$\Phi$ Thermodynamic Potential.
$\Psi$ Free energy.

$\alpha$ Expansion coefficient.
$\beta$ Pressure coefficient.
$\epsilon$ Individual or molecular energy.
$\theta$ Thermometric temperature.
$\kappa$ Ratio of the two specific heats.
$\lambda$ Wave-length, or heat of reaction.
$\mu$ Gibbs potential.
$\nu$ Frequency.

## Part V

## THEORY OF RELATIVITY

$A, B, C$ Arbitrary vector or tensor.
$E$ (**E**) Electric field-strength.
$F$ Electromagnetic field-tensor.
$G$ Cofactor in determinant of metric fundamental tensor.
$H$ (**H**) Magnetic field-strength.

$K$ (K) Force.
$R$ Riemann tensor or scalar (in § 197 radius of curvature of space).
$S$ Curvature tensor.
$T$ Matter tensor.

$b$ (b) Acceleration.
$c$ Velocity of light.
$ds$ World-line-element.
$g$ Metric fundamental tensor.
$m$ Mass.
$n$ Refractive index.
$q$ Minkowski-velocity.
$s$ Four-current.
$t$ Time.
$v$ Velocity.

$\alpha$ Gravitational radius (in § 182 dragging coefficient).
$\beta$ abbreviation for $1/\sqrt{1 - v^2/c^2}$.
$\delta$ Mixed fundamental tensor.
$\kappa$ Gravitational constant.
$\mu$ Proper mass.
$\nu$ Frequency.
$\varrho$ Density of mass or charge.
$\tau$ Proper time.
$d\tau$ Volume-element.
$\varphi, \psi$ Angle.

## THE UNIVERSAL CONSTANTS OF PHYSICS

| | |
|---|---|
| Elementary quantum of electricity | $4{\cdot}774 \times 10^{-10}$ gm.$^{\frac{1}{2}}$ cm.$^{\frac{3}{2}}$ sec.$^{-1}$ |
| Mass of an electron | $9{\cdot}00 \times 10^{-28}$ gm. |
| Mass of a proton | $1{\cdot}662 \times 10^{-24}$ gm. |
| Mass-ratio of proton of electron | 1846 |
| Loschmidt number | $6{\cdot}060 \times 10^{23}$ |
| Elementary quantum of action | $6{\cdot}545 \times 10^{-27}$ erg. sec. |
| Rydberg term | 109,737·1 cm.$^{-1}$ |
| Velocity of light | $2{\cdot}9980 \times 10^{10}$ cm. sec.$^{-1}$ |
| Boltzmann's constant | $1{\cdot}372 \times 10^{-16}$ erg. deg.$^{-1}$ |
| Gravitational constant | $6{\cdot}68 \times 10^{-8}$ gm.$^{-1}$ cm.$^3$ sec.$^{-2}$ |
| Mechanical equivalent of heat | $4{\cdot}186 \times 10^7$ erg. cal.$^{-1}$ |
| Gas constant | $8{\cdot}313 \times 10^7$ erg. deg.$^{-1}$ |
| ,, ,, (in thermal units) | 1·98 cal. deg.$^{-1}$ |

# ADDENDA AND CORRIGENDA

## Vol. I

§ 44, *second line before eqn.* (1) *: for* **K** *read* $\varrho\mathbf{K}$.
§ 46, *first line after eqn.* (6) *: for* **K** *read* $\varrho\mathbf{K}$.

## Vol. II

§ 127 : *add at end :—*

In 1928 a *theory of radioactivity* was developed from the principles of *wave-mechanics* simultaneously by *Gurney* and *Condon*, and by *Gamow*. This theory is related essentially to the principle of classical mechanics which excludes the motion of a particle from those regions in which its potential energy would be larger than its total constant energy (for then the kinetic energy would be negative and hence the velocity imaginary). According to the new theory of radioactivity, the atomic nuclei are surrounded by " walls " of high potential which, from the point of view of classical mechanics, represent insuperable obstacles for the constituents of the nucleus. From the point of view of wave-mechanics, however, there is always an assignable definite probability that, in an " attempt " on the potential wall, a particle contained in the nucleus will succeed in surmounting it. The wall owes its existence to the fact that, according to Coulomb's Law, which (as has been proved by manifold experiments) maintains an exact validity down to distances as small as $10^{-12}$ cm., the potential is inversely proportional to the distance, and therefore increases rapidly as the latter decreases : moreover, the repulsion experienced by an $\alpha$-particle indicates a positive potential. On the other hand, however, at distances of the order of magnitude of $10^{-13}$ cm. the potential must now be negative, because otherwise the $\alpha$-particles could not persist within the nucleus. Therefore, between $10^{-12}$ and $10^{-13}$ cm., the potential must exhibit an extremely steep negative gradient.

Considerations of wave-mechanics show that the prob-

ability of an escape past the potential wall is given by $e^{\gamma}$ where

$$\gamma = -\frac{4\pi}{h}\int \sqrt{2\,m(V-E)}\,dx.$$

In this expression $m$ denotes the mass of the particle, $E$ its energy, $V$ the potential energy, and $x$ the distance from the centre of the nucleus : the integral is to be taken over the "width of the wall," *i.e.*, over that interval of $x$ for which $V > E$. If the probability of an escape past the wall be multiplied by the revolution frequency of the $\alpha$-particle (about $10^{21}$), the disintegration constant is obtained. The same final value is found for this by various methods, *viz.*,

$$\lambda = \frac{1}{a}\sqrt{\frac{2E}{m}}\,e^{\gamma},$$

where $a$ denotes a length of the order of magnitude of the radius of the region enclosed by the wall. Exact calculations show that, in deriving the disintegration constant, it is necessary to take into account only that part of the potential course which obeys Coulomb's Law.

Since the disintegration constants depend chiefly on exponents in exponential expressions, a solution is afforded for the paradox that the half-value periods of the radio-elements differ by so many orders of magnitude, whereas the initial velocities of the $\alpha$-particles emitted by them are almost identical. The half-value periods derived on a wave-mechanical basis correspond in order of magnitude very closely to those experimentally found both for the short- and the long-lived elements. Geiger and Nuttall's relation can also be derived from the principles of wave-mechanics.

§ 133 : *for the last paragraph substitute :—*

Likewise, as Heisenberg has shown, a wave-mechanical resonance must also exist between the two similar atomic nuclei in elementary molecules, in consequence whereof either every second line in the bands must be completely missing, or else the lines exhibit an alternating intensity. The former is the case when no *spin* is associated with the nuclei ; the latter where a nuclear spin is present. In this way it has been found that there is no spin in the nuclei of *helium* and *oxygen*, whereas the existence of a spin in the nuclei of *hydrogen* and *nitrogen* was revealed. In point of fact, the number of nuclear electrons and protons is even for helium and oxygen, whereas there is an *odd* number of

electrons in the nitrogen nucleus, and the hydrogen nucleus is identical with a proton.

In 1929, *Giauque* and *Johnston* discovered, from the analysis of the *band-spectrum of oxygen*, the existence of molecules in which an oxygen atom of mass 16 is combined with one of mass 18 or of mass 17. The two *isotopes of oxygen* thus discovered are so rare that they could not have been detected by the mass-spectroscopic method of Aston. Of 10,000 oxygen atoms there are, perhaps, four of mass 18 and one of mass 17.

# NAME INDEX


Adams, 25, 148
Andrade, 172, 207
Andrews, 346
Anthony, 139
Aston, 17, 176, 182
Auger, 251
Avogadro, 333

Back, 130, 131
Balmer, 72
Barkla, 107
Barnett, 144
Bartels, 79
Beck, 219, 250
Becker, 174
Becquerel, 14
Bernoulli, 336
Berzelius, 236
Birge, 228
Black, 307, 324
Blackett, 213
Blake, 252
Bohr, 42, 73, 79, 101, 105, 113, 120, 143.
Boltzmann, 307–8, 338, 369
Born, 53, 57, 241, 253
Bothe, 256
Bowen, 90, 152
Boyle, 332
Brackett, 73
Bragg, W. H., 166 ff.
Bragg, W. L., 166 ff.
Bravais, 153
van den Broek, 37
Broéker, 181
de Broglie, L., 26, 41
de Broglie, M., 170
Brönsted, 181
Brown, 348
de Bruin, 101
Bunsen, 31, 33
Burger, 137

Cameron, 211–2
Campbell, 201
del Campo, 93
Cario, 150, 234
Carnot, 308
Catalán, 103
Celsuis, 304
Chadwick, 33, 214
Clapeyron, 308
Clausius, 290, 308, 327, 336
Cole, 174
Compton, 23, 256–7
Costa, 183
Coster, 35, 113, 120
Crookes, 12
Curie, M., 14, 201
Curie, P., 14, 142, 201
Curtis, 230

Dalton, 3, 9, 333
Darwin, 261
Davisson, 26, 172
Debierne, 201
Debye, 172, 241 ff., 358–9
Deluc, 324
Democritus, 3
Dempster, 177
Deslandres, 225
Dirac, 56, 117, 257 ff.
Dolejsek, 108
Dorgelo, 137, 151
Dorn, 184, 201
Draper, 371
Drude, 246
Duane, 252
Dulong, 307, 356
Dymond, 174

Eddington, 148, 152, 263
Ehrenhaft, 4, 8
Einsporn, 30, 92
Einstein, 22, 25, 144, 239, 349 ff., 357–8, 382 ff., 401, 410 ff., 440 ff.
Ellis, 207, 209
Elsasser, 174
Eötvös, 414
Epstein, 133
Eucken, 241
Euler, 43
Ewald, 241

Fajans, 202, 242
Falkenhagen, 243
Faraday, 10, 29, 244
Fermat, 42
Fermi, 245
Fitzgerald, 382
Fizeau, 395
Flamm, 83
Fortrat, 224
Fourier, 52, 307
Fowler, 79, 90, 100
Franck, 30, 92, 150–1, 233–4
Franz, 246
Fraunhofer, 129
Fresnel, 396
Friedrich, 165
Fuer, 251

Galileo, 388, 410
Gauss, 412 ff.
Gay-Lussac, 332
Gehrcke, 16
Geiger, 18, 203, 256
Gerlach, 139
Germer, 26, 172
Gibbs, R. C., 101
Gibbs, W., 308, 327 ff., 338
Giesel, 201
Goetze, 93
Goldstein, 12, 16, 101
Gordon, 257 ff.
Goudsmit, 61, 78,
Gray, 185
Grimm, 242
von Grosse, 196
Grüneisen, 361
Guldberg, 348
Günther, 174
Guth, 258

Haas, 45, 73, 264, 369
de Haas, 144
Haber, 240
Hahn, 201
Hallwachs, 16
Harkins, 181, 216 ff.
Heisenberg, 51 ff., 78, 94, 104, 117, 230, 239, 249
Heitler, 236
Helmholtz, 308, 312
Hertz, G., 29, 117
Hertz, H., 15
Herzfeld, 242
Hess, 186, 210-1
Heurlinger, 224
Hevesy, 35, 120, 181
Hittorf, 12

van't Hoff, 342
Holmes, 201
Hönigschmid, 176
Hopkins, 35
Houston, 73
Hubbard, 256
Hubble, 264, 448
Hückel, 242
Hund, 98, 103, 125, 143, 229, 242
Hunt, 252
Hylleraas, 94

Imes, 222
Ishida, 136

Jauncey, 258
Jeans, 262, 264, 296, 358
Jordan, 53, 78, 117
Joule, 308, 318, 333, 336

Kapitza, 131
Kármán, 241
Kaufmann, 12, 15
Kayser, 84
Kellner, 94
Kelvin, Lord, 304, 308, 318
Kirchhoff, 31, 33, 367
Kirsch, 214
Knipping, 165
Kolhörster, 211
Kossel, 111, 236, 342
Kovarik, 204, 207, 210
Kramers, 249
Kratzer, 224, 226
Krönig, 308, 336
Kuhn, 234
Kunsman, 173
Kurlbaum, 369

Ladenburg, 124, 249
Landé, 65 ff., 87
Langevin, 142
Lantsberry, 213
Laplace, 307
Laporte, 103–4
Larmor, 67
Laue, 153, 165
Lavoisier, 307
Lawson, 186, 201, 210
Lenard, 12, 16, 21
Lenz, 224, 229
Leverrier, 444
Lewis, 26, 228
Lindemann, 241, 357–8
London, 236 ff.
Long, 230
Lorentz, 125, 246, 382, 387, 405.

Loschmidt, 9, 338
Ludloff, 234
Lummer, 371
Lyman, 72, 94

McLennan, 100, 103, 125, 152
Madelung, 241, 358
Manley, 151
Marsden, 20, 213
Massieu, 312
Maupertuis, 43
Maxwell, 290, 308, 337
Mayer, 308, 335
Mecke, 228, 231
Meggers, 106
Meissner, 101, 150
Meitner, 201, 207, 210
Mendeléeff, 38
Merton, 181
Meyer, 38
Michelson, 378, 381
Millikan, 3 ff., 90, 211–2
Minkowski, 383–4
Moll, 137
Morley, 381
Moseley, 34, 108–9
Mulliken, 228

Nernst, 49, 241, 357–8, 361
Newton, 248, 382, 440
Nishina, 118
Noddack, 35
Nuttall, 203

Onnes, 247
Ornstein, 137

Paschen, 72, 79, 89, 101, 131
Pauli, 69, 117, 133
Perrin, 353
Petit, 307, 356
Pettersson, 214
Pfeffer, 342
Pickering, 79
Planck, 22, 370–2, 401
Poincaré, 383
Preston, 127
Prévost, 366
Pringsheim, 370
Prout, 215

Raman, 249
Ramsauer, 174
Ramsay, 37, 78, 185
Raoult, 342
Rayleigh, Lord, 44, 296, 358, 370
Regnault, 307
Reichenheim, 16
Richardson, 16
Riemann, 412, 433
Ritschl, 229
Ritz, 31, 72
Ross, 256
Royds, 185
Rubens, 358
Rubinowicz, 126
Rumford, 308
Runge, 84, 218
Rupp, 173–4
Russell, 62, 93, 104
Rutherford, 18, 28, 172, 184 ff., 201, 207, 212 ff.
Rydberg, 31, 84

Saha, 146
Sandemann, 230
Sänger, 244
Saunders, 62, 84, 93, 101
Scherrer, 172
Schrödinger, 47 ff., 86, 248, 259
Schwarzschild, 133, 440
Seemann, 169
Siegbahn, 108, 112
Simon, 256
de Sitter, 448
Slater, 78, 94, 104
Smekal, 113
Smith, 124
Smoluchowski, 308, 349, 354
Soddy, 78, 175, 184 ff., 201–2
Sommerfeld, 58, 63, 77, 103–4, 115, 133, 245, 252
Stark, 133
Stefan, 368
Stenström, 108
Stern, 139
Stewart, 242
Stokes, 23, 351
Stoner, 124
Stoney, 14
Sugiura, 237

Take, 35
Thibaud, 207
Thirring, 417
Thomson, G. P., 173, 177
Thomson, J. J., 8, 12, 17, 28, 176
Thomson, W., *vide* Kelvin, Lord
de la Tour, 346
Townsend, 8

Uhlenbeck, 61, 78

Waage, 348
van der Waals, 343
Walters, 103
Wang, 237
Weber, 356
Weiss, 143
Wentzel, 93, 113, 117, 250–1, 257
Westgren, 353
White, 101
Wiedemann, E., 149
Wiedemann, G., 246
Wien, M., 243
Wien, W., 150, 370–1
Wiener, 349
Wilke, 307
Wilson, C. T. R., 19
Wilson, W., 58
Wood, 149, 229
Wrede, 244

Zeeman, 125
Zisch, 240

# SUBJECT INDEX


Absolute zero,
  energy of, 49
  of temperature, 303
Absorption,
  band, 222
    head, 111
  coefficient, 204, 212
  spectrum, 33, 87
Actinium, 196, 201
  series, 197
Action, elementary quantum of, 22, 252, 372
Aggregation, state of, 321
Air, ionization of, 205
Alkali metals,
  arc spectra of, 84 ff.
  excitation and ionization potentials of, 89
  lowest spectral terms of, 87
  spark spectra of, 101
Alkaline earths,
  displaced terms of, 93
  spark spectra of, 89
Alpha
  -particles (rays), 15, 18, 27, 78, 185
    range of, 19, 202
    scattering of, 33
    scintillation of, 18, 213
  -transformation, 184
Alternation rule, Sommerfeld's, 104, 229
Analysis,
  positive ray, 17, 176
Angle, glancing, 161
Anode, 10
  rays, 16
Atoms, 3
  excitation of, 144 ff.
  hydrogen,
    angular momentum of, 42
    ionized, specific charge of, 12
    mass of, 12
  one-valency electron, spectrum of, 84 ff.
  stripped, spectra of, 90
  two-valency electron, spectrum of, 90 ff.
Avogadro's law, 33

Balmer's formula, 72
Band,
  absorption, 222
  branch, 224
  heads, law of, 225 ff.
  spectrum, 223 ff.
Beta
  -particles (rays), 15, 204–6
    secondary, 206
  -transformation, 184
Black body, 367
Boltzmann's
  constant, 307, 333
  entropy law, 307
  H-theorem, 338
Boyle's law, 332
de Broglie's theory, 41 ff.
Brownian movement, 348 ff.

Calcite, use of, in X-ray spectroscopy, 171
Calorie, 305
Calorimetry, 307
Canal rays, 16
Cathode, 10
  rays, 12 ff.
Centre, energy, 44
Charge number, nuclear, 33
Chemico-luminescence, 240
Chlorine, isotopy of, 177, 222
Coefficient,
  absorption, 204, 212, 367
  compressibility, 241, 313, 358
  diffusion, 351
  dragging, 395
  expansion, 313, 361
  pressure, 313
Combination rule,
  Laporte and Russell's, 104
  Ritz's, 31
Commutation rule, Heisenberg's, 53
Complexion, statistical, 268
Compound,
  heteropolar, 236
  homopolar, 236
Compressibility, coefficient of, 241, 313, 358
Compression, heat of, 316
Condition, Bohr's quantum, 42
Conductivity,
  electrical, 246

Conductivity—*continued.*
electrolytic, 243
thermal, 241, 246, 308
Cone, light-, 394
Constant,
Boltzmann's, 307, 333
chemical, 363
Curie, 142
dielectric, 244
disintegration, 186, 203, 480
equilibrium, 348
gas, 332
lattice, 154
Rydberg's, 73, 81
screening, 109
Contravariance, 421
Coordinates, Gaussian, 415
Copper, spectra of, 103
Covariance, 421
Critical state, 345
Crystal
powder, method of, 172
rotating, method of, 169
structure, 166 ff., 240 ff.
Curie
constant, 142
unit, 191
Current, four-, 406
Curve, saturation, 324

Dalton's law, 333
Density,
energy, 297
fluctuations of, 354
radiation, 366
Deposit, active, 195, 196, 198
Dew-point, 325
Diamagnetism, 141
Diamond,
atomic heat of, 357
structure of, 240
Dirac's equations, 259
Disintegration
constant, 186, 203, 480
theory, 184
Dispersion, 248 ff.
Displacement law, Wien's, 371
Dissociation,
heat of, 233
sensitized, 234
Distribution,
canonical, 273 ff., 338
modulus of, 274
Doublet,
hydrogen, 77
X-ray,
irregular, 117
regular, 115
Dragging coefficient, 395
Dulong and Petit's law, 356

Eclipse, solar (Einstein effect), 444
Effect,
Auger, 251
Compton, 23, 253 ff.
Doppler, 24
isotope, 222
Joule-Thomson, 318 ff.
packing, 182
Paschen-Back, 131
Peltier, 246
photo-electric, 15, 21, 250 ff.
Raman, 249
Stark, 133 ff., 139
Thomson, 246
Wien, 243
Zeeman, 14, 125 ff., 139
Efficiency (of cyclic process), 286
Electricity, elementary quantum of, 3 ff.
Electrolysis, 10
theory of, 242–3
Electrons, 14 ff., 174
angular momentum of, 59, 61
binding, 219
corresponding, 238
equivalent, 69, 95 ff.
free, 174, 245
grouping of, 122–3
impacts, 29, 151, 253
nuclear, 217
rays, 172, 252
recoil, 256
scattering of, 174
secondary, 210
selective reflection of, 26
specific charge of, 13, 83
spin of, 61, 68, 261
Elements,
disruption of, 212 ff.
long period, spectra of, 101 ff.
mixed, 176
pure, 177
sequence of, 33 ff.
short period, spectra of, 99 ff.
transformation of, 184 ff.
Emanation, 184, 191, 195–6, 201
Emulsion, 352
Energy,
centre, 44
density, 297
dissipation of, 307
dissociation, 231 ff.
elements, 22
excitation, 29, 145
free, 277, 289, 300, 310 ff., 360

Energy—*continued.*
level, 32
ionization, 30, 145, 151, 250
mass of, 23, 401
proper, 262, 401
Entropy, 306, 360 ff.
law, Boltzmann's, 307
Equilibrium,
constant, 348
radioactive, 187 ff.
redimentation, 352 ff.
statistical, 280, 297
thermodynamic, 307, 323, 366
Equipartition, theory of, 289
Equivalence,
chemical, 347
photo-chemical, law of, 239
principle of, 411
Excitation
energy, 29, 145
potential, 29
Excited atomic state, duration of, 149
Expansion, coefficient of, 313

Factor, resolving, 67, 128
Faraday's law, 10
Fermat's theorem, 42
Fermi's statistics, 245
Ferromagnetism, 143, 239
Field
metric, 416, 423
-potential, 405
tensor, electromagnetic, 405–7
Fine structure, 63, 261
Flames, conductivity of, 146
Fluctuational phenomena, molecular, 348
Fluorescence, 23, 148, 229, 251
extinction of, 149
sensitized, 150
Fluorine, valency of, 239
Force,
inertial, 409
Minkowski-, 397
relaxation, 243
statistical, 282
thermal, 311
Fortrat diagram, 225
Four
-current, 406
-potential, 405
Freedom, degree of, statistical, 282
Frequency,
characteristic, 249
nuclear vibrational, 232
Friction, internal, 337
Function, characteristic, 312

Gamma-rays, 15, 172, 207 ff.
secondary, 206, 210
Gas,
constant, 332
degeneration of, 245
perfect, 330 ff.
pressure, 336
Gay-Lussac's law, 333
Geodesic, 412, 426 ff.
Gram
-atom, 10
-molecule, 10
Graphite, structure of, 240
Gravitation, law of,
cosmological, 447
Einstein's, 413, 416, 435
Newton's, 440

H
-rays, 20, 213
-theorem, 338
Hafnium, discovery of, 120
Harkins' rules, 216 ff.
Head, absorption band, 111
Heat, 301
atomic, 355
capacity, 313
development of, by radium, 207
mechanical equivalent of, 305, 335
molar, 313
of reaction, 322
specific, 245, 314, 327, 334, 360
theorem, Nernst's, 359 ff.
Helium,
arc spectrum of, 94, 480
formation of, 185
ionized, spectrum of, 78
lines, displacement of, towards violet, 79 ff.
metastability of, 151
nucleus, 78
synthesis of, 182, 212, 263
Helmholtz's equation, 312
Humidity, 325
Hydrogen
atom,
constants of, 76
ionized, specific charge of, 12
mass of, 12
spectrum of, 72 ff., 261
doublet, 77
molecule,
dissociation of, sensitized, 234, 240
moment of inertia of, 231
nuclear interval of, 231
spectrum of, 480

Impacts, electron, 29, 151, 253
Inertia, moment of, molecular, 223, 231
Intercombinations, term, 91, 230
Interval,
  nuclear, 231
  rule, Landé's, 65, 106
Invariant, Riemann, 434
Iodine,
  fluorescence of, 229
  moment of inertia of, molecular, 231
Ionium, 195
Ionization
  energy, 30, 145, 151, 250
  of air, 205
  potential, 30 ff., 89
  spectrometer, 170
Ions, 10
  radii of, 242
Irreversibility, 308–10
Isobares, 180
Isotherms, 344
Isotope
  effect, 222
  figure, 218
Isotopes, separation of, 180–1
Isotopy, 175 ff., 229

Lattice,
  atomic, 240 ff., 358
  constant, 154
  cubic, 176–7
  line-, 154
  plane-, 155
  planes, 159 ff.
  space-, 156
Laue photograph, 165
Lead, atomic weight of, 200
Level, energy, 32, 209
Life-period, average, 187
Light
  -cone, 394
  -quanta, 22, 111
  rays, deflection of, 443–4
Lines,
  auroral, 152
  geodesic, 412, 426 ff.
  helium, displacement of, towards violet, 79 ff.
  resonance, 88, 92
  spectral, 31
    displacement of,
      (Doppler effect), 24
      towards red, 25, 445
    forbidden, 150–2
    intensities of, 137 ff.
  world-, 384
  X-ray, scheme of, 110, 119
Lorentz-transformation, 387
Loschmidt number, 9–11, 186, 351–3, 372
Luminescence,
  chemico-, 240
  thermal, 144

Mache-unit, 191
Magneton, 140, 143
Mass
  action, law of, 346 ff.
  formula, relativistic, 398
  inertial and gravitational, 413–4
  -spectroscopy, 176
  stationary, 398
Matrix calculus, 53
Matter,
  dissolution of into light, 261 ff.
  tensor, 407, 416, 434
  wave theory of, 26, 41 ff.
Maupertuis's theorem, 43
Mean free path, 336
Mechanics,
  of $q$-numbers, Dirac's, 56
  quantum, 51 ff.
Melting-point, constancy of, 324
Membrane, semipermeable, 339
Mercury
  (element), arc spectrum of, 92
  (planet), orbit of, 444
Mesothorium, 198
Metastability, atomic, 150–1
Michelson's experiment, 378
Minerals, age of, 200
Minkowski
  -force, 397
  -velocity, 397
  -world, 384, 412
Model, atomic, Rutherford's, 28
Modulus values, equalization of, 281
Molecules,
  diameter of, 338
  polar, 243
  velocity of, 336
Momentum, angular, quantized, 42
Motion, perihelion, 443
Multiple proportions, law of, 9
Multiplicity, spectroscopic, 63 ff., 239

Nebulium, 152
Neon,
  isotopy of, 176
  metastability of, 150
Nernst's heat theorem, 359 ff.
Nitrogen,
  molecular spectrum of, 480
  nucleus, disruption of, 213

Nucleus, atomic, 27
  disruption of, 212 ff.
  motion of, 79
Number,
  Loschmidt, 9–11, 186, 351–3, 372
  nuclear charge (atomic), 33
  quantum, *vide* quantum number
  wave-, 32

Observation, individual, method of, 4
Opalescence, optical, 354
Operator, mathematical, 54, 258
Orbits, quantized, 42
Orthohelium, 94
Oscillator, linear, 48
Oxygen,
  isotopy of, 481
  molecular spectrum of, 481
  valency of, 239

Paramagnetism, 124, 141
Parhelium, 94
Pauli's principle, 69–70, 124, 219, 237
Perihelion motion, 443
Period (of elements), 38
  half-value, 186, 480
Phase, 321
  -rule, 330
Photochemistry, 234, 239
Photons, 26
Piezo-electricity, 242
Planck's law of radiation, 370
Point, world-, 384
Polonium, 196, 201
Potassium, radioactivity of, 201
Potential,
  contact, 246
  equivalent, 16, 20
  excitation, 29 ff., 89, 110
  field-, 405
  four-, 405
  Gibbs, 328
  ionization, 30 ff., 89
  thermodynamic, 312, 329
  transformation, 151
Power, emissive, 367
Pressure, 291
  coefficient of, 313
  gas, 336
  osmotic, 339 ff.
  partial, law of, 333
  saturation, 324
Preston's rule, 127
Probability, statistical, 268 ff., 279, 289
Process,
  adiabatic, 284
  cyclic, 284 ff.
    efficiency of, 286
Process—*continued*
  isomodular, 284
  isothermal, 311
  reversible, 276, 279
Protoactinium, 196, 201
Protons, 28, 33, 218
Pyro-electricity, 242

Q-numbers, mechanics of, 56
Quantum
  condition, Bohr's, 42
  light-, 22, 31, 111
  mechanics, 51 ff.
  number,
    effective, 86
    inner, 63
    magnetic, 65
    principal, 61, 69, 86
    subordinate, 61, 69, 77, 85
  of action, elementary, 22, 252, 372
  of electricity, elementary, 3 ff.
  statistics, 245

Radiation,
  cavity, 366 ff.
  density, 366
  law of, 370
  nuclear, 208
  resonance, 148, 249
Radioactivity, 14, 184 ff.
  wave-mechanical theory of, 479
Radioelement, 193 ff.
Radiothorium, 198
Radium, 14, 184–5, 187, 195, 201
  development of heat by, 207
Radius, gravitational, 441
Rare
  earths, 40, 105, 120
  gases, 37
    arc spectra of, 101
Ratio, packing, 183
Rayleigh and Jeans's principle, 296
Rayleigh's radiation law, 370
Rays,
  $\alpha$-, *vide* alpha-rays
  anode, 16
  atomic, 139
  Becquerel, 14
  $\beta$-, *vide* beta-rays
  canal, 16
  cathode, 12 ff.
  cosmic, 211–2
  electron, 172, 252
  $\gamma$-, *vide* gamma-rays
  H-, 20, 213
  recoil, 206
  residual, 241, 358
  X-, *vide* X-rays

Reaction, heat of, 322
Recoil
  electrons, 256
  rays, 206
Refraction, double, 242
Relativity, principle of, 383
Resonance
  line, 88, 92
  quantum-mechanical, 94, 236, 239, 480
  radiation, 148, 249
Ritz correction, 86
Rock-salt, structure of, 167
Rotating crystal, Braggs' method of, 169
Rotator, 50
Rubidium, radioactivity of, 201
Runge's rule, 128
Rydberg's constant, 73, 81
  correction, 86

Scalar,
  Riemann, 416
  Schrödinger's field-, 46, 57, 248
Schrödinger's,
  equation, 46, 56, 236
  field-scalar, 46, 57, 248
  theory, 46 ff.
Scintillation (of α-particles), 18, 213
Selection rule, Rubinowicz's, 126
Series,
  actinium, 197
  Bergmann, 84
  fundamental, 84
  principal, 84
  spectral, 32
  subordinate, 84
  thorium, 199
  transformation, 193 ff.
  uranium-radium, 194
  X-ray, 108 ff.
Silicon, arc spectrum of, 100
Simultaneity, 389
Sirius, companion of,
  density of, 148
  displacement of spectral lines of, 25
Sky, colour of, 354
Solution, colloidal, 352
Space, spherical, 446
Spectrometer, ionization, 170
Spectroscopy,
  mass-, 176 ff.
  solar, 146
  X-ray, 34, 108 ff.
Spectrum,
  absorption, 33, 87
  arc,
    of mercury, 92
    of silicon, 100
Spectrum—*continued*
  band, 223 ff.
  helium, ionized, 78, 230
  hydrogen,
    atomic, 72 ff., 261
    molecular, 229
  magnetic, 204 ff.
  many-line, 230
  molecular, 220 ff.
  rotation, 223
  spark, 89, 101, 117
  vibration, rotation-, 221
Stars, internal temperature of, 147
State,
  critical, 345
  distribution of, 268
  equation of, 282, 343 ff.
  quantities of, 282
    reduced, 346
Stefan's law, 368
Stokes's rule, 23, 149
Summation rule, Ornstein's, 137
Sun, age of, 263
  spectrum of, 146
Supraconductivity, 247
Susceptibility, magnetic, 142
Suspension, 352
Symbol, three-indices, 429
System,
  heterogeneous, 327 ff.
  periodic, 38–40, 118 ff., 231
  proper, 396
  statistically homogeneous, 282

Temperature, 303
  absolute, 303, 332
  characteristic, 355
  saturation, 324
  thermodynamic, 320
  zero of, absolute, 303
Tension, vapour, 325
Tensor,
  contraction of, 423
  covariant derivative of, 432
  curvature, 416, 434
  field, electromagnetic, 405–7
  matter, 407, 416, 434
  metric fundamental, 415, 423 ff.
  mixed, 432
  Riemann, 433
  skew-symmetrical, 402
  world-, 401 ff.
    vector divergence of, 403
Term,
  displaced, 93
  fundamental, 32
    rule, Slater's, 104
  multiplet, 63 ff., 239
    regular, 92

Term—*continued*.
spectral, 31, 60, 85
X-ray, 113
Thermodynamics, two principal laws of, 306, 308
Thermometry, 304, 307
Thermoscope, 302
Thorium, 198
series, 199
Three-indices symbol, 429
Time,
concept of, 382
proper, 396
relaxation, 243
Transformation,
$\alpha$-, 184
$\beta$-, 184
Lorentz-, 387
potential, 151
series, 193 ff.
Triple-point, 327

Universe,
extent of, 264, 445 ff.
radius of curvature of, 447
Uranium, 190, 193
lead, 200
-radium series, 194

Valency, 238
Vapour.
superheated, 325
tension, 325
Vector, world-, 384, 393
rotation of, 403
Velocities,
composition of, 391–3
group, 43–5
molecular, 336
Maxwell's theorem of distribution of, 290
Velocity, Minkowski-, 397
Vibration, characteristic, 293, 295
Virial, 290–1

Wave,
group, 43
material, 26, 42 ff., 250, 363
–number, 32
Weights, atomic, 9, 33, 175
Wiedemann and Franz's law, 245
Wien's displacement law, 371
World,
-line, 384
Minkowski-, 384, 412
-point, 384
-tensor, 401 ff.
vector divergence of, 403
-vector, 384, 393
rotation of, 403

X-rays, 153, 252
characteristic, 107, 206
doublet,
irregular, 117
regular, 115
interference of, 165 ff.
lines, scheme of, 110, 114
spark spectra, 117
spectroscopy, 108 ff.
terms, 113
wave-lengths of, 169 ff.
X-unit, 171

PRINTED IN GREAT BRITAIN BY THE WHITEFRIARS PRESS, LTD.,
LONDON AND TONBRIDGE.

530.151 H11E TV5 1928 V>

a39001 006961307b